Algebraic theory of regular languages

Libor Polák *
Department of Mathematics, Masaryk University
Janáčkovo nám. 2a, 662 95 Brno, Czech Republic
http://www.math.muni.cz/~polak

*Supported by the Min. of Education of the Czech Republic under the project MSM 143100009

Finite automata

A **nondeterministic finite automata** (NFA) :
$\mathcal{A} = (Q, A, E, I, T)$ where Q is a finite non-empty set of states, A is a finite non-empty alphabet, $E \subseteq Q \times A \times Q$ is a set of transitions, $I, T \subseteq Q$ are the sets of (all) initial and terminal states.

Finite automata

A **nondeterministic finite automata** (NFA) :
$\mathcal{A} = (Q, A, E, I, T)$ where Q is a finite non-empty set of states, A is a finite non-empty alphabet, $E \subseteq Q \times A \times Q$ is a set of transitions, $I, T \subseteq Q$ are the sets of (all) initial and terminal states.

$\mathcal{A}$ is **deterministic** (DFA), if $I = \{i\}$ and
$(p, a, q),\ (p, a, r) \in E \implies q = r$.

Finite automata

A **nondeterministic finite automata** (NFA) :
$\mathcal{A} = (Q, A, E, I, T)$ where Q is a finite non-empty set of states, A is a finite non-empty alphabet, $E \subseteq Q \times A \times Q$ is a set of transitions, $I, T \subseteq Q$ are the sets of (all) initial and terminal states.

$\mathcal{A}$ is **deterministic** (DFA), if $I = \{i\}$ and
$(p, a, q),\ (p, a, r) \in E \implies q = r$.

A **complete DFA** : $\mathcal{A} = (Q, A, \cdot, i, T)$ where $\cdot : Q \times A \to Q$. It naturally extends to $\cdot : Q \times A^* \to Q$.

Finite automata

A **nondeterministic finite automata** (NFA) :
$\mathcal{A} = (Q, A, E, I, T)$ where Q is a finite non-empty set of states, A is a finite non-empty alphabet, $E \subseteq Q \times A \times Q$ is a set of transitions, $I, T \subseteq Q$ are the sets of (all) initial and terminal states.

$\mathcal{A}$ is **deterministic** (DFA), if $I = \{i\}$ and
$(p, a, q),\ (p, a, r) \in E \implies q = r$.

A **complete DFA** : $\mathcal{A} = (Q, A, \cdot, i, T)$ where $\cdot : Q \times A \to Q$. It naturally extends to $\cdot : Q \times A^* \to Q$.

A **language** over alphabet A : $L \subseteq A^*$ (sometimes $L \subseteq A^+$).

Successful path :

$(q_1, a_1, q_2),\ (q_2, a_2, q_3), \ldots, (q_k, a_k, q_{k+1}),\ k \geq 0,\ q_1 \in I,\ q_{k+1} \in T$.

Its **label** : $a_1 \ldots\ a_k$.

Successful path :
$(q_1, a_1, q_2),\ (q_2, a_2, q_3), \ldots, (q_k, a_k, q_{k+1}),\ k \geq 0,\ q_1 \in I,\ q_{k+1} \in T$.
Its **label** : $a_1 \ldots\ a_k$.

A language **accepted** by an automaton $\mathcal{A}$ is the set of all labels of successful paths; notation $L(\mathcal{A})$.

Successful path :
$(q_1, a_1, q_2),\ (q_2, a_2, q_3), \ldots, (q_k, a_k, q_{k+1}),\ k \geq 0,\ q_1 \in I,\ q_{k+1} \in T.$
Its **label** : $a_1 \ldots\ a_k$.

A language **accepted** by an automaton $\mathcal{A}$ is the set of all labels of successful paths; notation $L(\mathcal{A})$.

Minimalization of a complete DFA $\mathcal{A}$: remove all non-reachable states and factorize Q through μ where
$p\mu q \ \Leftrightarrow\ L(Q, A, \cdot, p, T) = L(Q, A, \cdot, q, T).$

Successful path :
$(q_1, a_1, q_2),\ (q_2, a_2, q_3), \ldots, (q_k, a_k, q_{k+1}),\ k \geq 0,\ q_1 \in I,\ q_{k+1} \in T.$
Its **label** : $a_1 \ldots\ a_k$.

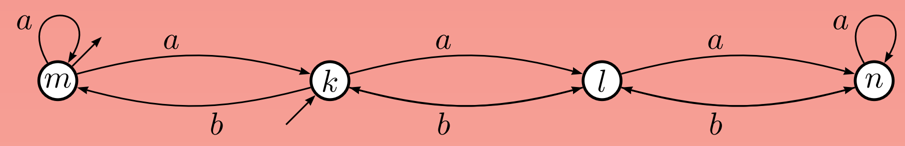
A language **accepted** by an automaton $\mathcal{A}$ is the set of all labels of successful paths; notation $L(\mathcal{A})$.

Minimalization of a complete DFA $\mathcal{A}$: remove all non-reachable states and factorize Q through μ where
$p \mu q \ \Leftrightarrow\ L(Q, A, \cdot, p, T) = L(Q, A, \cdot, q, T)$.

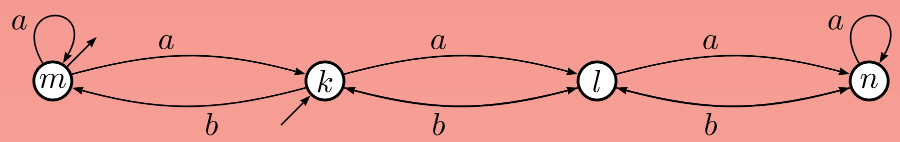
It can be implemented in $O(n \lg n)$ time, where $n = |Q|$; later we will see that such a local procedure is also a global one.

Theorem State minimalization of a NFA is a PSPACE-complete problem.

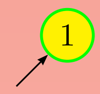
Determinization of an NFA



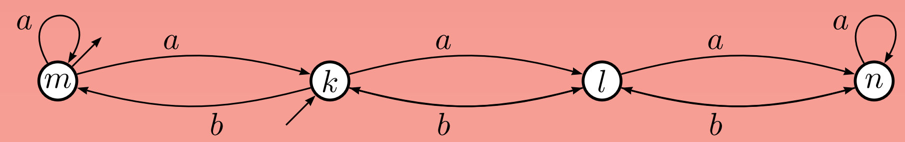
Determinization of an NFA



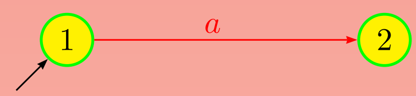
$$1 = \{k\},\ 2 = \{l\},\ 3 = \{l, m\},\ 4 = \{n\},\ 5 = \{k, n\}, \ldots$$



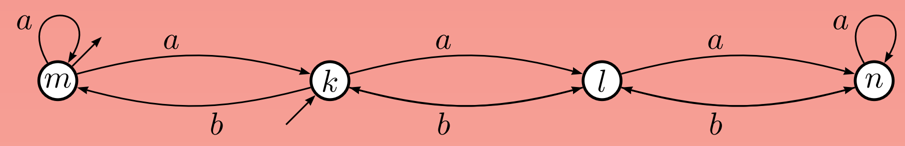
Determinization of an NFA



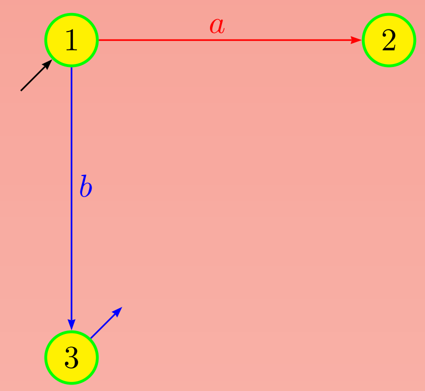
$$1 = \{k\},\ 2 = \{l\},\ 3 = \{l, m\},\ 4 = \{n\},\ 5 = \{k, n\}, \ldots$$



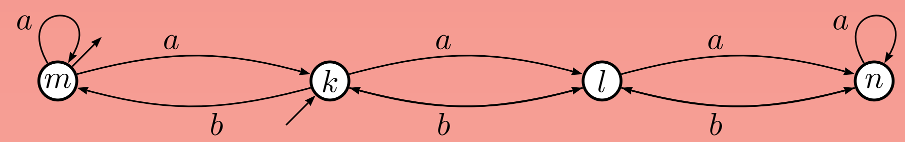
Determinization of an NFA



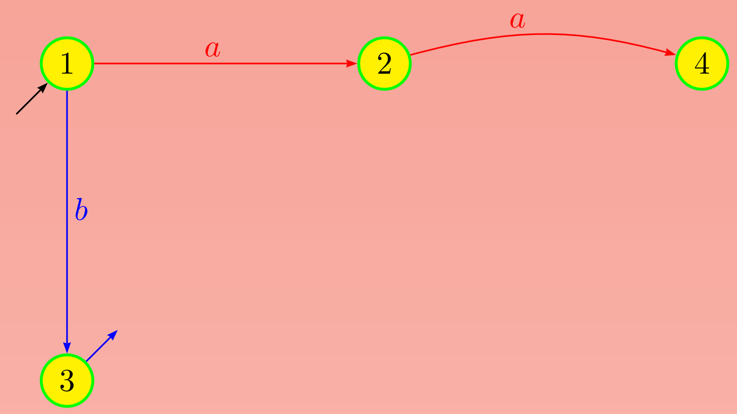
$$1 = \{k\},\ 2 = \{l\},\ 3 = \{l, m\},\ 4 = \{n\},\ 5 = \{k, n\}, \ldots$$



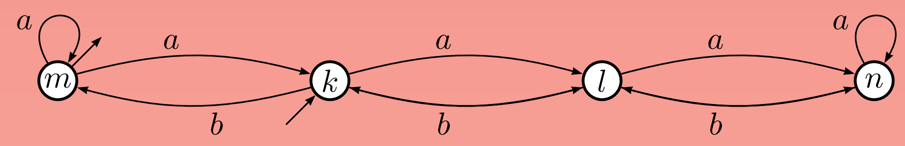
Determinization of an NFA



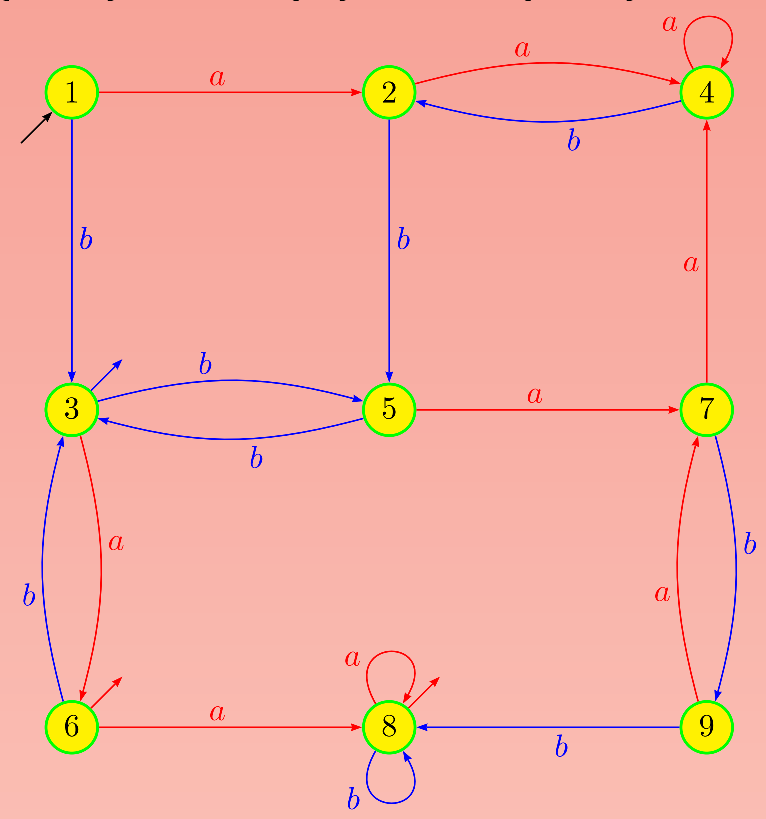
$$1 = \{k\},\ 2 = \{l\},\ 3 = \{l, m\},\ 4 = \{n\},\ 5 = \{k, n\}, \ldots$$



Determinization of an NFA



$$1 = \{k\},\ 2 = \{l\},\ 3 = \{l, m\},\ 4 = \{n\},\ 5 = \{k, n\}, \ldots$$



Theorem For each NFA there exists a complete DFA accepting the same language. Consequently, the class of all languages accepted by finite automata is closed with respect to complements.

Theorem For each NFA there exists a complete DFA accepting the same language. Consequently, the class of all languages accepted by finite automata is closed with respect to complements.

Syntactic structures

An **idempotent semiring** is a structure $(S, \cdot, \vee)$ where
$(S, \cdot)$ is a monoid with the neutral element 1,
$(S, \vee)$ is a semilattice with the smallest element 0,
$(\ \forall\ a, b, c \in S\ )(\ a(b \vee c) = ab \vee ac$ and $(a \vee b)c = ac \vee bc\ )$,
and $(\ \forall\ a \in S\ )\ a0 = 0a = 0$.

Theorem For each NFA there exists a complete DFA accepting the same language. Consequently, the class of all languages accepted by finite automata is closed with respect to complements.

Syntactic structures

An **idempotent semiring** is a structure $(S, \cdot, \vee)$ where
$(S, \cdot)$ is a monoid with the neutral element 1,
$(S, \vee)$ is a semilattice with the smallest element 0,
$(\ \forall\ a, b, c \in S\ )(\ a(b \vee c) = ab \vee ac$ and $(a \vee b)c = ac \vee bc\ )$,
and $(\ \forall\ a \in S\ )\ a0 = 0a = 0$.
Such a structure becomes an ordered monoid with respect to the relation $\leq$ defined by $a \leq b \ \Leftrightarrow\ a \vee b = b,\ a, b \in S$.

Theorem For each NFA there exists a complete DFA accepting the same language. Consequently, the class of all languages accepted by finite automata is closed with respect to complements.

Syntactic structures

An **idempotent semiring** is a structure $(S, \cdot, \vee)$ where
$(S, \cdot)$ is a monoid with the neutral element 1,
$(S, \vee)$ is a semilattice with the smallest element 0,
$(\ \forall\ a, b, c \in S\ )(\ a(b \vee c) = ab \vee ac$ and $(a \vee b)c = ac \vee bc\ )$,
and $(\ \forall\ a \in S\ )\ a0 = 0a = 0$.
Such a structure becomes an ordered monoid with respect to the relation $\leq$ defined by $a \leq b \ \Leftrightarrow\ a \vee b = b,\ a, b \in S$.
Let $A^{\square}$ denote the set of all finite subsets of A^*.

Note that this set with the operations $U \cdot V = \{ \, uv \mid u \in U, \, v \in V \, \}$ and usual union form a free idempotent semiring over the set A.

Note that this set with the operations $U \cdot V = \{\ uv \mid u \in U,\ v \in V\ \}$ and usual union form a free idempotent semiring over the set A.
Let $L \subseteq A^*$ be a regular language. We define the following relations :
for $u, v, u_1, \ldots, u_k, v_1, \ldots, v_l \in A^*$,

$$u \ \approx_L \ v \text{ if and only if } (\ \forall\ x, y \in A^*\ )(\ xuy \in L \ \Leftrightarrow\ xvy \in L\ )\ ,$$

Note that this set with the operations $U \cdot V = \{\ uv \mid u \in U,\ v \in V\ \}$ and usual union form a free idempotent semiring over the set A.
Let $L \subseteq A^*$ be a regular language. We define the following relations :
for $u, v, u_1, \ldots, u_k, v_1, \ldots, v_l \in A^*$,

$$u \approx_L v \text{ if and only if } (\ \forall\ x, y \in A^*\ )(\ xuy \in L \ \Leftrightarrow\ xvy \in L\ )\ ,$$

$\{u_1, \ldots, u_k\} \sim_L \{v_1, \ldots, v_l\}$ if and only if

$$(\ \forall\ x, y \in A^*\ )(\ xu_1y, \ldots, xu_ky \in L \ \Leftrightarrow\ xv_1y, \ldots, xv_ly \in L\ )\ .$$

Note that this set with the operations $U \cdot V = \{ \ uv \mid u \in U, \ v \in V \ \}$ and usual union form a free idempotent semiring over the set A. Let $L \subseteq A^*$ be a regular language. We define the following relations :
for $u, v, u_1, \ldots, u_k, v_1, \ldots, v_l \in A^*$,

$$u \approx_L v \text{ if and only if } (\ \forall\ x, y \in A^*\ )(\ xuy \in L \ \Leftrightarrow \ xvy \in L\ )\ ,$$

$\{u_1, \ldots, u_k\} \ \sim_L \ \{v_1, \ldots, v_l\}$ if and only if

$$(\ \forall\ x, y \in A^*\ )(\ xu_1y, \ldots, xu_ky \in L \ \Leftrightarrow \ xv_1y, \ldots, xv_ly \in L\ )\ .$$

The factor-structures $(\mathsf{O}\,(L), \cdot, 1_{\approx_L}) = (A^*, \cdot, 1)/\approx_L$ and $(\mathsf{S}\,(L), \cdot, \vee) = (A^{\Box}, \cdot, \cup)/\sim_L$ are called the **syntactic monoid** and the **syntactic semiring** of the language L.

the **syntactic semiring** of the language L. The assignments

$$\kappa_L : u \mapsto u \approx_L \text{ and } \phi_L : \{u_1, \ldots, u_k\} \mapsto \{u_1, \ldots, u_k\} \sim_L$$

are called **syntactic monoid/semiring homomorphisms**.

the **syntactic semiring** of the language L. The assignments

$$\kappa_L : u \mapsto u \approx_L \text{ and } \phi_L : \{u_1, \dots, u_k\} \mapsto \{u_1, \dots, u_k\} \sim_L$$

are called **syntactic monoid/semiring homomorphisms**.

The monoid $(\mathsf{O}(L), \cdot)$ is ordered by the relation

$$v \approx_L \leq u \approx_L \quad \text{iff} \quad (\forall\, x, y \in A^*)\,(\, xuy \in L \Rightarrow xvy \in L\,)$$

and we speak about the **ordered syntactic monoid**.

the **syntactic semiring** of the language L. The assignments

$$\kappa_L : u \mapsto u \approx_L \text{ and } \phi_L : \{u_1, \ldots, u_k\} \mapsto \{u_1, \ldots, u_k\} \sim_L$$

are called **syntactic monoid/semiring homomorphisms**.

The monoid $(\mathsf{O}\,(L), \cdot)$ is ordered by the relation

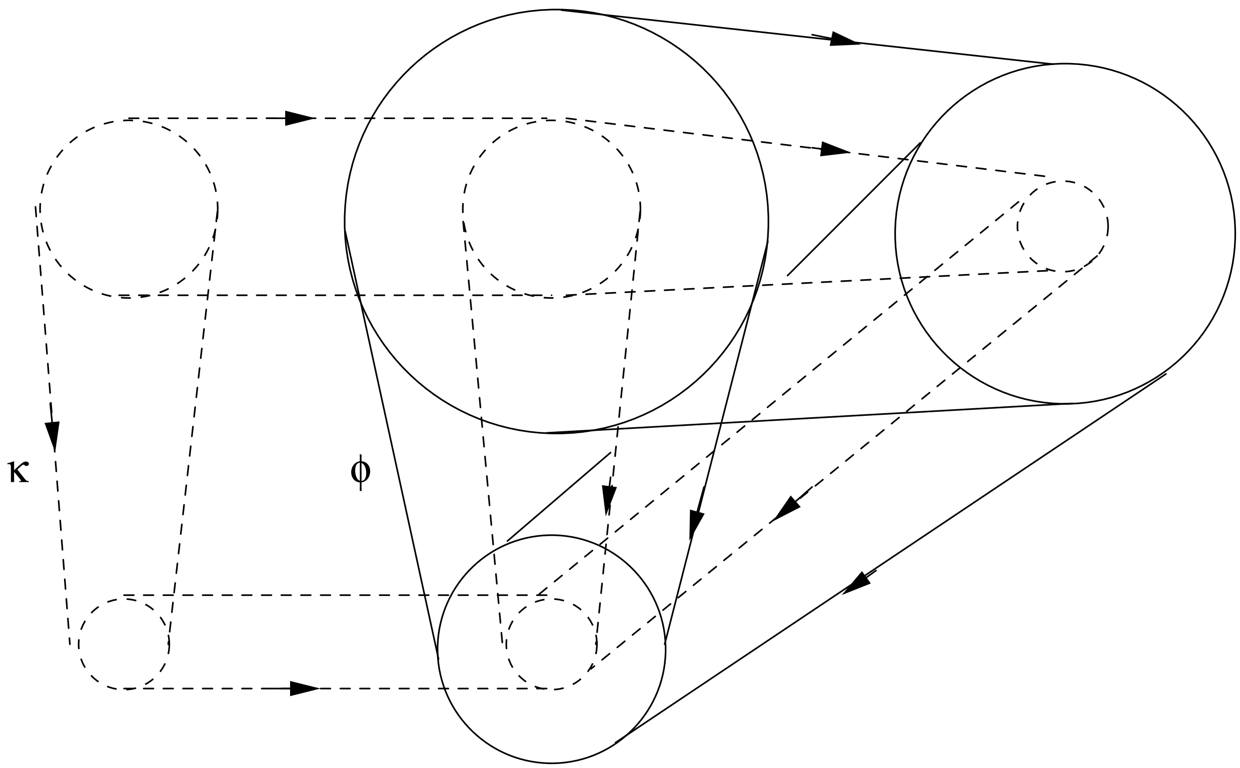
$$v \approx_L \leq \; u \approx_L \quad \text{iff} \;\; (\, \forall\, x, y \in A^* \,)\; (\, xuy \in L \;\Rightarrow\; xvy \in L \,)$$

and we speak about the **ordered syntactic monoid**.

The structures $(\mathsf{O}\,, \cdot, \leq)$ and $(\mathsf{S}\,, \cdot, \vee)$ are equationally independent; certain relations are explained by the following diagram.

A^* $\qquad$ $A^\square$ $\qquad$ $\mathsf{H}\,(\mathsf{O}\,,\leq)$ (all hereditary subsets)

$$u \;\mapsto\; \{u\}\;, \quad \{u_1,\ldots,u_k\} \;\mapsto\; (\; u_1\approx,\ldots,u_k\approx\;]$$



$\mathsf{O} = A^*/\approx$ $\qquad$ $\mathsf{S} = A^\square/\sim$

$$u\approx \;\mapsto\; \{u\}\sim$$

Regular expressions

Let A be a fixed finite alphabet, let $X = \{x_1, \ldots, x_m\}$ $(m \geq 0)$ be a finite set of variables. We denote by Reg (A, X) the set of all terms in the language of nullary operational symbols $0, a_1, \ldots, a_n$, binary operational symbols $\cdot, +$ and unary operational symbol * over the variables $x_1, \ldots, x_m$. The elements of Reg (A, X) are called **regular expressions** over A in X.

Regular expressions

Let A be a fixed finite alphabet, let $X = \{x_1, \ldots, x_m\}$ $(m \geq 0)$ be a finite set of variables. We denote by Reg (A, X) the set of all terms in the language of nullary operational symbols $0, a_1, \ldots, a_n$, binary operational symbols $\cdot, +$ and unary operational symbol * over the variables $x_1, \ldots, x_m$. The elements of Reg (A, X) are called **regular expressions** over A in X.

Let $\overline{r}$ be the realization of $r \in$ Reg (A, X) in the algebra $(2^{A^*}, \emptyset, \{a_1\}, \ldots, \{a_n\}, \cdot, \cup, ^*)$ of all subsets of A^* where $\cdot$ is the catenation, $\cup$ is the set-theoretical union and * is the Kleene star.

$K \cdot L = \{ \ uv \mid u \in K, \ v \in L \ \}$,
$K^* = \{ \ u_1 \ldots u_k \mid k \geq 0, \ u_1, \ldots, u_k \in A^* \ \}$

Adding to the signature binary $\cap$ and unary ^{c}, realized by intersection and complement, we get **generalized** regular expressions.

Adding to the signature binary $\cap$ and unary ^{c}, realized by intersection and complement, we get **generalized** regular expressions.

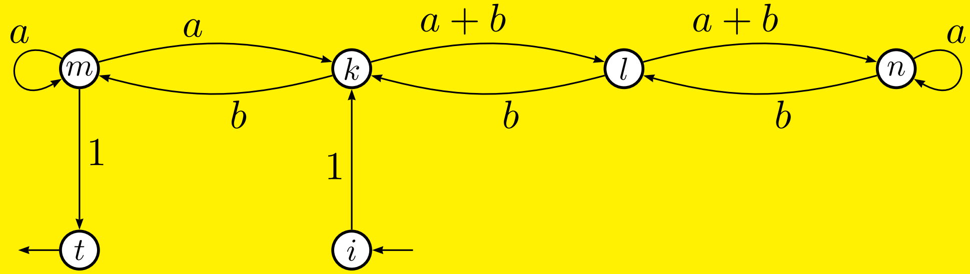
We will use $(m > 0)$-ary expressions later when solving equations. So we fix now $X = \emptyset,\ m = 0$.

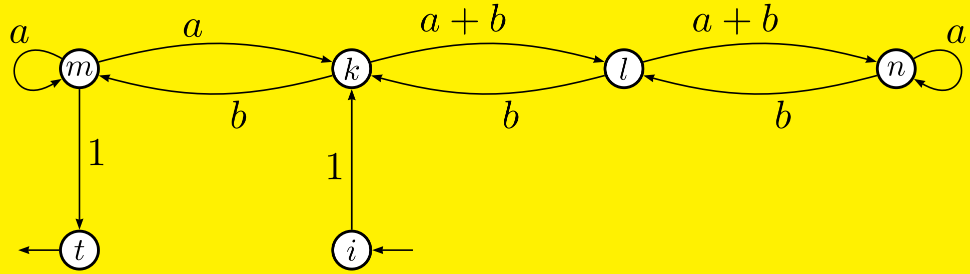
Adding to the signature binary $\cap$ and unary ^{c}, realized by intersection and complement, we get **generalized** regular expressions.

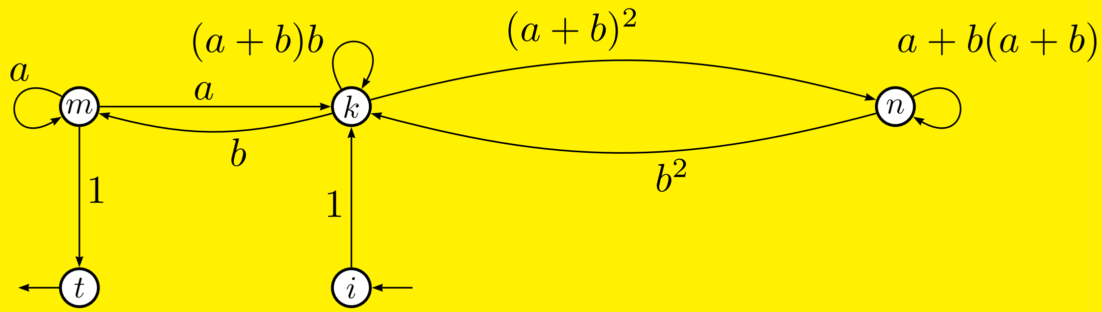
We will use $(m > 0)$-ary expressions later when solving equations. So we fix now $X = \emptyset,\ m = 0$.

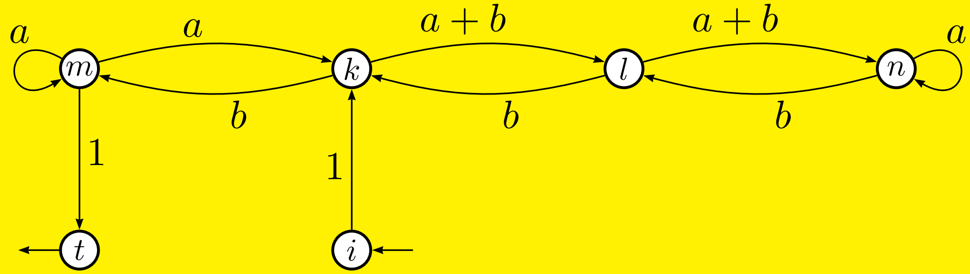
Theorem (Kleene)
The languares realized by regular expressions are exactly the languages accepted by NFA's. We speak about **regular** languages.

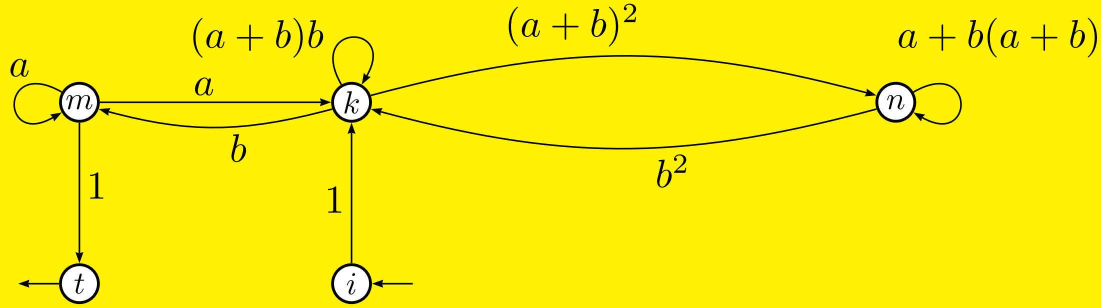
Proof $\Leftarrow$: McNaughton - Yamada algorithm :

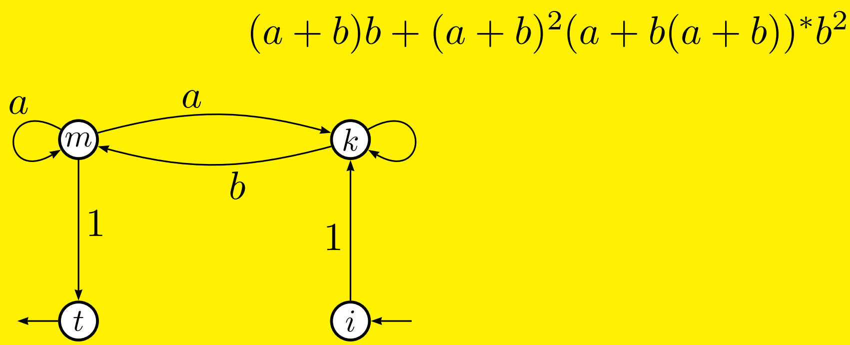

a
a
a + b
a + b
a
m
k
l
n
b
b
b
1
1
t
i


a
a
a + b
a + b
a
m
k
l
n
b
b
b
1
1
t
i

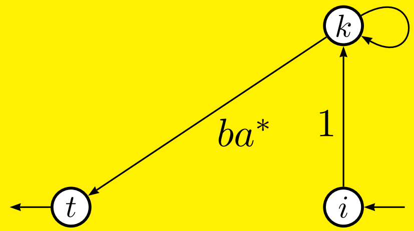

(a + b)b
(a + b)^2
a + b(a + b)
a
a
m
k
n
b
b^2
1
1
t
i


a
a
$a + b$
$a + b$
a
m
k
l
n
b
b
b
1
1
t
i

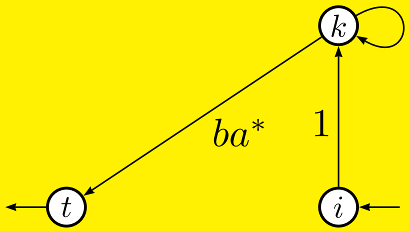

$(a+b)b$
$(a+b)^2$
$a + b(a+b)$
a
a
m
k
n
b
b^2
1
1
t
i


$(a+b)b + (a+b)^2(a+b(a+b))^*b^2$
a
a
m
k
b
1
1
t
i

$$(a+b)b+(a+b)^2(a+b(a+b))^*b^2+ba^*a$$

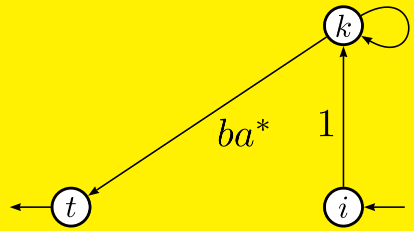


$$(a+b)b+(a+b)^2(a+b(a+b))^*b^2+ba^*a$$



Result : $((a+b)b+(a+b)^2(a+b(a+b))^*b^2+ba^*a)^*ba^*$.

$$(a+b)b+(a+b)^2(a+b(a+b))^*b^2+ba^*a$$



Result : $((a+b)b+(a+b)^2(a+b(a+b))^*b^2+ba^*a)^*ba^*$.

$\Rightarrow$: ...

Star height of generalized regular expression r = the maximal number of nested stars in r.

Star height of a language L = the minimal star height of a regular expression realizing L.

Star height of generalized regular expression $r =$ the maximal number of nested stars in r.

Star height of a language $L =$ the minimal star height of a regular expression realizing L.

We define the generalized star height of L in a similar way.

Star height of generalized regular expression $r =$ the maximal number of nested stars in r.

Star height of a language $L =$ the minimal star height of a regular expression realizing L.

We define the generalized star height of L in a similar way.

Example Let $A = \{a, b\}$. The language
$L_n = \{\ u \in A^* \mid |u|_a \equiv |u|_b \bmod 2^n\ \}$ is of star height n.

Star height of generalized regular expression r = the maximal number of nested stars in r.

Star height of a language L = the minimal star height of a regular expression realizing L.

We define the generalized star height of L in a similar way.

Example Let $A = \{a, b\}$. The language
$L_n = \{\ u \in A^* \mid |u|_a \equiv |u|_b \bmod 2^n\ \}$ is of star height n.

Big problem Compute the star height of a given regular language.

Star height of generalized regular expression $r =$ the maximal number of nested stars in r.

Star height of a language $L =$ the minimal star height of a regular expression realizing L.

We define the generalized star height of L in a similar way.

Example Let $A = \{a, b\}$. The language
$L_n = \{ \ u \in A^* \mid |u|_a \equiv |u|_b \bmod 2^n \ \}$ is of star height n.

Big problem Compute the star height of a given regular language.

Even bigger problem Is there a regular language of generalized star height at least two ?

Let $G = (V, E)$ be a finite oriented graph. We define the **loop complexity** $\mathsf{lc}G$ of G inductively by

$\mathsf{lc}(\{v\}, \emptyset) = 0$, $\mathsf{lc}(\{v\}, \{(v, v)\}) = 1$ and for $|V| \geq 2$

for strongly connected G : $\mathsf{lc}G = \min\{\, \mathsf{lc}(G \setminus v) \mid v \in V\,\} + 1$,

otherwise : $\mathsf{lc}G = \max\{\, \mathsf{lc}H \mid H$ is a s.c. component of $G\,\}$.

Let $G = (V, E)$ be a finite oriented graph. We define the **loop complexity** lc$G$ of G inductively by

lc$(\{v\}, \emptyset) = 0$, lc$(\{v\}, \{(v, v)\}) = 1$ and for $|V| \geq 2$

for strongly connected G : lc$G = \min\{$ lc$(G \setminus v) \mid v \in V \} + 1$,

otherwise : lc$G = \max\{$ lc$H \mid H$ is a s.c. component of $G \}$.

Problem What is a complexity of computing the loop complexity ?

Let $G = (V, E)$ be a finite oriented graph. We define the **loop complexity** lc$G$ of G inductively by

lc$(\{v\}, \emptyset) = 0$, lc$(\{v\}, \{(v, v)\}) = 1$ and for $|V| \geq 2$

for strongly connected G : lc$G = \min\{$ lc$(G \setminus v) \mid v \in V\} + 1$,

otherwise : lc$G = \max\{$ lc$H \mid H$ is a s.c. component of $G\}$.

Problem What is a complexity of computing the loop complexity ?

Theorem (Eggan)
For a NFA $\mathcal{A}$ the number lc$\mathcal{A}$ is equal to the minimum of the star heights of expressions we get from $\mathcal{A}$ running the McNaughton-Yamada algorithm.

Schützenberger, Simon,...

Theorem (Schützenberger)
A language $L \subseteq A^*$ is a star-free (= can be presented by a generalized regular expression without $*$) if and only if the syntactic monoid of L is aperiodic (= it contains only trivial subgroups).

Schützenberger, Simon,...

Theorem (Schützenberger)
A language $L \subseteq A^*$ is a star-free ($=$ can be presented by a generalized regular expression without $*$) if and only if the syntactic monoid of L is aperiodic ($=$ it contains only trivial subgroups).

Homework Show that $(ab + ba)^*$ defines over $A = \{a, b\}$ a star-free language.

Schützenberger, Simon,...

Theorem (Schützenberger)
A language $L \subseteq A^*$ is a star-free (= can be presented by a generalized regular expression without $*$) if and only if the syntactic monoid of L is aperiodic (= it contains only trivial subgroups).

Homework Show that $(ab + ba)^*$ defines over $A = \{a, b\}$ a star-free language. Hint : consider first $(ab)^*$.

Schützenberger, Simon,...

Theorem (Schützenberger)
A language $L \subseteq A^*$ is a star-free ($=$ can be presented by a generalized regular expression without $*$) if and only if the syntactic monoid of L is aperiodic ($=$ it contains only trivial subgroups).

Homework Show that $(ab+ba)^*$ defines over $A=\{a,b\}$ a star-free language. Hint : consider first $(ab)^*$.

Solution $A^* = \emptyset^{\mathsf{c}},\ (ab)^* = 1 + aA^* \cap A^*b \cap (A^*aaA^*)^{\mathsf{c}} \cap (A^*bbA^*)^{\mathsf{c}},$

Schützenberger, Simon,...

Theorem (Schützenberger)
A language $L \subseteq A^*$ is a star-free (= can be presented by a generalized regular expression without $*$) if and only if the syntactic monoid of L is aperiodic (= it contains only trivial subgroups).

Homework Show that $(ab+ba)^*$ defines over $A = \{a, b\}$ a star-free language. Hint : consider first $(ab)^*$.

Solution $A^* = \emptyset^c$, $(ab)^* = 1 + aA^* \cap A^*b \cap (A^*aaA^*)^c \cap (A^*bbA^*)^c$,

$$(ab+ba)^* = ((1+A^*b)b(ab)^*(1+bA^*)+(1+A^*a)(ab)^*a(1+aA^*))^c.$$

Theorem (Simon)
A language $L \subseteq A^*$ is a finite union of finite intersections of the languages $A^*a_1A^*a_2 \dots A^*a_kA^*$ and their complements if and only if its syntactic monoid is $\mathcal{J}$-trivial.

$a\mathcal{J}b$ in M if $MaM = MbM$

Theorem (Simon)
A language $L \subseteq A^*$ is a finite union of finite intersections of the languages $A^*a_1A^*a_2 \ldots A^*a_kA^*$ and their complements if and only if its syntactic monoid is $\mathcal{J}$-trivial.

$a\mathcal{J}b$ in M if $MaM = MbM$

Theorem
A language $L \subseteq A^*$ is a finite union (of finite intersections) of the languages $A^*a_1A^*a_2 \ldots A^*a_kA^*$ if and only if its ordered syntactic monoid satisfies $x \leq 1$.

Theorem (Simon)
A language $L \subseteq A^*$ is a finite union of finite intersections of the languages $A^*a_1A^*a_2 \ldots A^*a_kA^*$ and their complements if and only if its syntactic monoid is $\mathcal{J}$-trivial.

$a\mathcal{J}b$ in M if $MaM = MbM$

Theorem
A language $L \subseteq A^*$ is a finite union (of finite intersections) of the languages $A^*a_1A^*a_2 \ldots A^*a_kA^*$ if and only if its ordered syntactic monoid satisfies $x \leq 1$.

Example
$A^*aA^*aA^*bA^* \cap A^*bA^*aA^* =$
$A^*aA^*aA^*bA^*aA^* \cup A^*aA^*bA^*aA^*bA^* \cup A^*bA^*aA^*aA^*bA^*$.

subword : chlast $\leq$ chlapskáslast

subword : chlast $\leq$ chlapskáslast (How many times ?)

subword : chlast $\leq$ chlapskáslast (How many times ?)

Lemma (Higman 1952, proof by Conway with the main idea by Nash-Williams 1963)
Each infinite set of words over a finite alphabet A contains a pair of different words u, v such that $u \leq v$.

subword : chlast $\leq$ chlapskáslast (How many times ?)

Lemma (Higman 1952, proof by Conway with the main idea by Nash-Williams 1963)
Each infinite set of words over a finite alphabet A contains a pair of different words u, v such that $u \leq v$.

(Lothaire, Combinatorics on words, Algebraic combinatorics on words, Applied combinatorics on words.)

subword : chlast $\leq$ chlapskáslast (How many times ?)

Lemma (Higman 1952, proof by Conway with the main idea by Nash-Williams 1963)
Each infinite set of words over a finite alphabet A contains a pair of different words u, v such that $u \leq v$.

(Lothaire, Combinatorics on words, Algebraic combinatorics on words, Applied combinatorics on words.)

Proof Suppose that $u_1, u_2, \ldots$ is an infinite sequence of words such that $i < j \;\Rightarrow\; u_i \not\leq u_j$.

subword : chlast $\leq$ chlapskáslast (How many times ?)

Lemma (Higman 1952, proof by Conway with the main idea by Nash-Williams 1963)
Each infinite set of words over a finite alphabet A contains a pair of different words u, v such that $u \leq v$.

(Lothaire, Combinatorics on words, Algebraic combinatorics on words, Applied combinatorics on words.)

Proof Suppose that $u_1, u_2, \ldots$ is an infinite sequence of words such that $i < j \;\Rightarrow\; u_i \not\leq u_j$.
Consider such sequences with minimal $|u_1|$ (the **lenght** of u_1).
Among them consider sequences with minimal $|u_2|$,...

subword : chlast $\leq$ chlapskáslast (How many times ?)

Lemma (Higman 1952, proof by Conway with the main idea by Nash-Williams 1963)
Each infinite set of words over a finite alphabet A contains a pair of different words u, v such that $u \leq v$.

(Lothaire, Combinatorics on words, Algebraic combinatorics on words, Applied combinatorics on words.)

Proof Suppose that $u_1, u_2, \ldots$ is an infinite sequence of words such that $i < j \;\Rightarrow\; u_i \not\leq u_j$.
Consider such sequences with minimal $|u_1|$ (the **lenght** of u_1).
Among them consider sequences with minimal $|u_2|$,...
There is $a \in A$ such that infinitely many from $u_1, u_2, \ldots$ start

with a.

with a. So let $u_{i_1} = av_{i_1}, u_{i_2} = av_{i_2}, \ldots,\ i_1 < i_2 < \ldots$.

with a. So let $u_{i_1} = av_{i_1}, u_{i_2} = av_{i_2}, \ldots,\ i_1 < i_2 < \ldots.$

Taking $u_1, u_2, \ldots, u_{i_1-1}, v_{i_1}, v_{i_2}, \ldots$ we get a contradiction with the "minimality". $\square$

with a. So let $u_{i_1} = av_{i_1}, u_{i_2} = av_{i_2}, \ldots,\ i_1 < i_2 < \ldots$.

Taking $u_1, u_2, \ldots, u_{i_1-1}, v_{i_1}, v_{i_2}, \ldots$ we get a contradiction with the "minimality". □

Proof of the Theorem

$\Rightarrow$: Let $L = A^*a_1A^*a_2\ldots A^*a_kA^*$, that is L is the set of all words having $a_1 \ldots a_k$ as a subword.

with a. So let $u_{i_1} = av_{i_1}, u_{i_2} = av_{i_2}, \ldots,\ i_1 < i_2 < \ldots$.

Taking $u_1, u_2, \ldots, u_{i_1 - 1}, v_{i_1}, v_{i_2}, \ldots$ we get a contradiction with the "minimality". $\square$

Proof of the Theorem

$\Rightarrow$: Let $L = A^* a_1 A^* a_2 \ldots A^* a_k A^*$, that is L is the set of all words having $a_1 \ldots a_k$ as a subword. Clearly

$(\forall\ p, q, r \in A^*)\ (\ pq \in L \implies puq \in L\ ) \quad (*)$.

with a. So let $u_{i_1} = av_{i_1}, u_{i_2} = av_{i_2}, \ldots,\ i_1 < i_2 < \ldots$.

Taking $u_1, u_2, \ldots, u_{i_1-1}, v_{i_1}, v_{i_2}, \ldots$ we get a contradiction with the "minimality". $\square$

Proof of the Theorem

$\Rightarrow$: Let $L = A^*a_1A^*a_2 \ldots A^*a_kA^*$, that is L is the set of all words having $a_1 \ldots a_k$ as a subword. Clearly
$(\forall\ p, q, r \in A^*) (pq \in L \implies puq \in L) \quad (*)$.
The same is true for finite unions of such languages.

with a. So let $u_{i_1} = av_{i_1}, u_{i_2} = av_{i_2}, \ldots,\ i_1 < i_2 < \ldots$.

Taking $u_1, u_2, \ldots, u_{i_1 - 1}, v_{i_1}, v_{i_2}, \ldots$ we get a contradiction with the "minimality". $\square$

Proof of the Theorem

$\Rightarrow$: Let $L = A^* a_1 A^* a_2 \ldots A^* a_k A^*$, that is L is the set of all words having $a_1 \ldots a_k$ as a subword. Clearly
$(\forall\ p, q, r \in A^*) \ (pq \in L \implies puq \in L) \quad (*)$.
The same is true for finite unions of such languages.

$\Leftarrow$: Let a language satisfy the condition $(*)$. Let M be the set of all minimal words from L with respect to $\leq$.

with a. So let $u_{i_1} = av_{i_1}, u_{i_2} = av_{i_2}, \ldots,\ i_1 < i_2 < \ldots$.

Taking $u_1, u_2, \ldots, u_{i_1-1}, v_{i_1}, v_{i_2}, \ldots$ we get a contradiction with the "minimality". $\square$

Proof of the Theorem

$\Rightarrow$: Let $L = A^*a_1A^*a_2\ldots A^*a_kA^*$, that is L is the set of all words having $a_1\ldots a_k$ as a subword. Clearly
$(\ \forall\ p, q, r \in A^*\ )\ (\ pq \in L \implies puq \in L\ ) \quad (*)$.
The same is true for finite unions of such languages.

$\Leftarrow$: Let a language satisfy the condition $(*)$. Let M be the set of all minimal words from L with respect to $\leq$.
We have $M = \{\ a_{1,1}\ldots a_{1,l_1}, \ldots, a_{k,1}\ldots a_{k,l_k}\}$.

with a. So let $u_{i_1} = av_{i_1}, u_{i_2} = av_{i_2}, \ldots,\ i_1 < i_2 < \ldots$.

Taking $u_1, u_2, \ldots, u_{i_1-1}, v_{i_1}, v_{i_2}, \ldots$ we get a contradiction with the "minimality". $\Box$

Proof of the Theorem

$\Rightarrow$: Let $L = A^*a_1A^*a_2 \ldots A^*a_kA^*$, that is L is the set of all words having $a_1 \ldots a_k$ as a subword. Clearly
$(\ \forall\ p, q, r \in A^*\ )\ (\ pq \in L \implies puq \in L\ ) \quad (*)$.
The same is true for finite unions of such languages.

$\Leftarrow$: Let a language satisfy the condition $(*)$. Let M be the set of all minimal words from L with respect to $\leq$.
We have $M = \{\ a_{1,1} \ldots a_{1,l_1}, \ldots, a_{k,1} \ldots a_{k,l_k}\}$.
Put $L = A^*a_{1,1}A^* \ldots A^*a_{1,l_1}A^* \ldots$. $\Box$

Canonical minimal DFA

For $u \in A^*$ and $K \subseteq A^*$ we set $u^{-1}K = \{ v \in A^* \mid uv \in K \}$. Let $L \subseteq A^*$ and put $\mathsf{U} = \{ u^{-1}L \mid u \in A^* \}$, $\mathcal{D} = (\mathsf{D}, A, \cdot, L, T)$ where
$u^{-1}L \cdot a = a^{-1}(u^{-1}L)$ $(= (ua)^{-1}L)$,
$u^{-1}L \in T$ iff $1 \in u^{-1}L$ (iff $u \in L$).

Canonical minimal DFA

For $u \in A^*$ and $K \subseteq A^*$ we set $u^{-1}K = \{ \ v \in A^* \mid uv \in K \ \}$. Let $L \subseteq A^*$ and put $\mathsf{U} = \{ \ u^{-1}L \mid u \in A^* \ \}$, $\mathcal{D} = (\mathsf{D}\,, A, \cdot, L, T)$ where $u^{-1}L \cdot a = a^{-1}(u^{-1}L)$ $(= (ua)^{-1}L)$,
$u^{-1}L \in T$ iff $1 \in u^{-1}L$ (iff $u \in L$).

We see immediately that $\mathcal{D}$ accepts L : $L \xrightarrow{u} u^{-1}L \in T$ iff $u \in L$.

Canonical minimal DFA

For $u \in A^*$ and $K \subseteq A^*$ we set $u^{-1}K = \{ v \in A^* \mid uv \in K \}$. Let $L \subseteq A^*$ and put U $= \{ u^{-1}L \mid u \in A^* \}$, $\mathcal{D} = (\text{D}\,, A, \cdot, L, T)$ where
$u^{-1}L \cdot a = a^{-1}(u^{-1}L)$ $(= (ua)^{-1}L)$,
$u^{-1}L \in T$ iff $1 \in u^{-1}L$ (iff $u \in L$).

We see immediately that $\mathcal{D}$ accepts L : $L \xrightarrow{u} u^{-1}L \in T$ iff $u \in L$.

Proposition (Brzozowski)
Let $\mathcal{A} = (Q, A, \cdot, i, F)$ be a DFA with all states reachable accepting the language $L \subseteq A^*$. Then the mapping $\phi : i \cdot u \mapsto u^{-1}L$ is a surjective automata homomorphism. Moreover, if $\mathcal{A}$ is minimal, then ϕ is an isomorphism.

Canonical minimal DFA

For $u \in A^*$ and $K \subseteq A^*$ we set $u^{-1}K = \{ v \in A^* \mid uv \in K \}$. Let $L \subseteq A^*$ and put $\mathsf{U} = \{ u^{-1}L \mid u \in A^* \}$, $\mathcal{D} = (\mathsf{D}, A, \cdot, L, T)$ where
$u^{-1}L \cdot a = a^{-1}(u^{-1}L)$ $(= (ua)^{-1}L)$,
$u^{-1}L \in T$ iff $1 \in u^{-1}L$ (iff $u \in L$).

We see immediately that $\mathcal{D}$ accepts L : $L \xrightarrow{u} u^{-1}L \in T$ iff $u \in L$.

Proposition (Brzozowski)
Let $\mathcal{A} = (Q, A, \cdot, i, F)$ be a DFA with all states reachable accepting the language $L \subseteq A^*$. Then the mapping $\phi : i \cdot u \mapsto u^{-1}L$ is a surjective automata homomorphism. Moreover, if $\mathcal{A}$ is minimal, then ϕ is an isomorphism.

Proof Correctness, that is, $i \cdot u = i \cdot v$ gives $u^{-1}L = v^{-1}L$: Indeed,

$w \in u^{-1}L$ yields $uw \in L,\ vw \in L$ and $w \in v^{-1}L$; similarly $\supseteq$.

$w \in u^{-1}L$ yields $uw \in L,\ vw \in L$ and $w \in v^{-1}L$; similarly $\supseteq$.

Surjectivity : a preimage of $u^{-1}L$ is $i \cdot u$.

$w \in u^{-1}L$ yields $uw \in L,\ vw \in L$ and $w \in v^{-1}L$; similarly $\supseteq$.

Surjectivity : a preimage of $u^{-1}L$ is $i \cdot u$.

Moreover, $\phi : i \mapsto L$ and $i \cdot u \in F$ iff $u \in L$ iff $u^{-1}L \in T$.

$w \in u^{-1}L$ yields $uw \in L,\ vw \in L$ and $w \in v^{-1}L$; similarly $\supseteq$.

Surjectivity : a preimage of $u^{-1}L$ is $i \cdot u$.

Moreover, $\phi : i \mapsto L$ and $i \cdot u \in F$ iff $u \in L$ iff $u^{-1}L \in T$.

Finally we show that

$$u^{-1}L = v^{-1}L \text{ gives } L(Q, A.\cdot, i \cdot u, F) = L(Q, A.\cdot, i \cdot v, F)\,.$$

$w \in u^{-1}L$ yields $uw \in L,\ vw \in L$ and $w \in v^{-1}L$; similarly $\supseteq$.

Surjectivity : a preimage of $u^{-1}L$ is $i \cdot u$.

Moreover, $\phi : i \mapsto L$ and $i \cdot u \in F$ iff $u \in L$ iff $u^{-1}L \in T$.

Finally we show that

$$u^{-1}L = v^{-1}L \text{ gives } L(Q, A.\cdot, i \cdot u, F) = L(Q, A.\cdot, i \cdot v, F)\,.$$

Indeed, from $i \cdot uw \in F$ we consecutively get
$uw \in L,\ w \in u^{-1}L,\ w \in v^{-1}L,\ vw \in L,\ i \cdot vw \in F$; $\supseteq$ is similar.

Constructions of syntactic structures

Theorem The transformation monoid of a minimal complete DFA $\mathcal{A} = (Q, a, \cdot, i, T)$ is isomorphic to the syntactic monoid $(\mathrm{O}\,(L), \cdot)$ of the language $L = L(\mathcal{A})$. An isomorphism is given by $(\ [u] : q \mapsto q \cdot u,\ q \in Q\ ) \mapsto u \approx,\ u \in A^*$.

Constructions of syntactic structures

Theorem The transformation monoid of a minimal complete DFA $\mathcal{A} = (Q, a, \cdot, i, T)$ is isomorphic to the syntactic monoid $(\mathsf{O}\,(L), \cdot)$ of the language $L = L(\mathcal{A})$. An isomorphism is given by $(\ [u] : q \mapsto q \cdot u,\ q \in Q\ ) \mapsto u \approx,\ u \in A^*$.

Proof We show that $u \approx v \iff [u] = [v]$.

Constructions of syntactic structures

Theorem The transformation monoid of a minimal complete DFA $\mathcal{A} = (Q, a, \cdot, i, T)$ is isomorphic to the syntactic monoid $(\mathsf{O}\,(L), \cdot)$ of the language $L = L(\mathcal{A})$. An isomorphism is given by $(\ [u] : q \mapsto q \cdot u,\ q \in Q\ ) \mapsto u \approx,\ u \in A^*$.

Proof We show that $u \approx v \iff [u] = [v]$. Indeed,

$u \approx v$ iff
$(\ \forall\ p, q \in A^*\ )\ puq \in L\ \Leftrightarrow\ pvq \in L$, that is iff

Constructions of syntactic structures

Theorem The transformation monoid of a minimal complete DFA $\mathcal{A} = (Q, a, \cdot, i, T)$ is isomorphic to the syntactic monoid $(\mathrm{O}\,(L), \cdot)$ of the language $L = L(\mathcal{A})$. An isomorphism is given by $(\ [u] : q \mapsto q \cdot u,\ \ q \in Q\ \ ) \mapsto u \approx,\ \ u \in A^*$.

Proof We show that $u \approx v \iff [u] = [v]$. Indeed,

$u \approx v$ iff
$(\ \forall\ p, q \in A^*\ )\ puq \in L\ \Leftrightarrow\ pvq \in L$, that is iff
$(\ \forall\ p, q \in A^*\ )\ i \cdot puq \in T\ \Leftrightarrow\ i \cdot pvq \in T$, that is iff

Constructions of syntactic structures

Theorem The transformation monoid of a minimal complete DFA $\mathcal{A} = (Q, a, \cdot, i, T)$ is isomorphic to the syntactic monoid $(\mathrm{O}\,(L), \cdot)$ of the language $L = L(\mathcal{A})$. An isomorphism is given by
$(\ [u] : q \mapsto q \cdot u,\ \ q \in Q\ ) \mapsto u \approx,\ \ u \in A^*$.

Proof We show that $u \approx v \iff [u] = [v]$. Indeed,

$u \approx v$ iff
$(\ \forall\ p, q \in A^*\ )\ puq \in L\ \ \Leftrightarrow\ \ pvq \in L$, that is iff
$(\ \forall\ p, q \in A^*\ )\ i \cdot puq \in T\ \ \Leftrightarrow\ \ i \cdot pvq \in T$, that is iff
$(\ \forall\ p \in A^*\ )\ i \cdot pu = i \cdot pv$ (minimality), that is iff

Constructions of syntactic structures

Theorem The transformation monoid of a minimal complete DFA $\mathcal{A} = (Q, a, \cdot, i, T)$ is isomorphic to the syntactic monoid $(\mathrm{O}\,(L), \cdot)$ of the language $L = L(\mathcal{A})$. An isomorphism is given by $(\ [u] : q \mapsto q \cdot u,\ \ q \in Q\ ) \mapsto u \approx,\ \ u \in A^*$.

Proof We show that $u \approx v \iff [u] = [v]$. Indeed,

$u \approx v$ iff
$(\ \forall\ p, q \in A^*\ )\ puq \in L\ \Leftrightarrow\ pvq \in L$, that is iff
$(\ \forall\ p, q \in A^*\ )\ i \cdot puq \in T\ \Leftrightarrow\ i \cdot pvq \in T$, that is iff
$(\ \forall\ p \in A^*\ )\ i \cdot pu = i \cdot pv$ (minimality), that is iff
$(\ \forall\ p \in A^*\ )\ (i \cdot p)[u] = (i \cdot p)[v]$ that is iff

Constructions of syntactic structures

Theorem The transformation monoid of a minimal complete DFA $\mathcal{A} = (Q, a, \cdot, i, T)$ is isomorphic to the syntactic monoid $(\mathsf{O}\,(L), \cdot)$ of the language $L = L(\mathcal{A})$. An isomorphism is given by $(\ [u] : q \mapsto q \cdot u,\ q \in Q\ ) \mapsto u \approx,\ u \in A^*$.

Proof We show that $u \approx v \iff [u] = [v]$. Indeed,

$u \approx v$ iff
$(\ \forall\ p, q \in A^*\ )\ puq \in L\ \Leftrightarrow\ pvq \in L$, that is iff
$(\ \forall\ p, q \in A^*\ )\ i \cdot puq \in T\ \Leftrightarrow\ i \cdot pvq \in T$, that is iff
$(\ \forall\ p \in A^*\ )\ i \cdot pu = i \cdot pv$ (minimality), that is iff
$(\ \forall\ p \in A^*\ )\ (i \cdot p)[u] = (i \cdot p)[v]$ that is iff
$(\ \forall\ r \in Q\ )\ r[u] = r[v]$ (each state is reachable from i), that is iff

$$[u] = [v].$$

$[u] = [v]$.

Clearly, $[u] \circ [v] = [uv] \mapsto (uv) \approx = u \approx \ \cdot v \approx$. $\Box$

$[u] = [v]$.

Clearly, $[u] \circ [v] = [uv] \mapsto (uv) \approx\ = u \approx\ \cdot\ v \approx .$ $\square$

Often we take the canonical minimal DFA $\mathcal{D}$ in place of (a general) $\mathcal{A}$.

$[u] = [v]$.

Clearly, $[u] \circ [v] = [uv] \mapsto (uv) \approx = u \approx \ \cdot v \approx$. $\Box$

Often we take the canonical minimal DFA $\mathcal{D}$ in place of (a general) $\mathcal{A}$.

Let U $= \{ \ u_1^{-1}L \cap \cdots \cap u_k^{-1}L \mid k \geq 0, \ u_1, \ldots, u_k \in A^* \ \}$.

$[u] = [v]$.

Clearly, $[u] \circ [v] = [uv] \mapsto (uv) \approx = u \approx \ \cdot v \approx$. $\square$

Often we take the canonical minimal DFA $\mathcal{D}$ in place of (a general) $\mathcal{A}$.

Let U $= \{ \ u_1^{-1}L \cap \cdots \cap u_k^{-1}L \mid k \geq 0, \ u_1, \ldots, u_k \in A^* \ \}$.

The action of a letter $a \in A$ is now given by

$$(q_1 \cap \cdots \cap q_m) \cdot a = q_1 \cdot a \ \cap \cdots \cap \ q_m \cdot a \ .$$

It can be extended to transitions induced by finite sets of words by

$$q \cdot \{u_1, ..., u_k\} = q \cdot u_1 \cap \cdots \cap q \cdot u_k \text{ for } q \in \mathsf{U}, \ u_1, ..., u_k \in A^* \ .$$

Again, denote by $[u_1, \ldots, u_k]$ the transformation of U given by the set $\{u_1, \ldots, u_k\}$.

Again, denote by $[u_1, \ldots, u_k]$ the transformation of U given by the set $\{u_1, \ldots, u_k\}$.

Theorem The transformation monoid $\{\ [u_1, \ldots, u_k] \mid k \geq 0,\ u_1, \ldots, u_k \in A^*\ \}$ with the operations of composition and $[u_1, \ldots, u_k] \vee [v_1, \ldots, v_l] = [u_1, \ldots, u_k, v_1, \ldots, v_l]$. is isomorphic to the syntactic semiring $(\mathsf{S}\,(L), \cdot, \vee)$ of the language L. An isomorphism is given by: $[u_1, \ldots, u_k] \mapsto \{u_1, \ldots, u_k\}\sim$.

Again, denote by $[u_1, \ldots, u_k]$ the transformation of U given by the set $\{u_1, \ldots, u_k\}$.

Theorem The transformation monoid $\{\ [u_1, \ldots, u_k] \mid k \geq 0,\ u_1, \ldots, u_k \in A^*\ \}$ with the operations of composition and $[u_1, \ldots, u_k] \vee [v_1, \ldots, v_l] = [u_1, \ldots, u_k, v_1, \ldots, v_l]$. is isomorphic to the syntactic semiring $(\mathsf{S}\,(L), \cdot, \vee)$ of the language L. An isomorphism is given by: $[u_1, \ldots, u_k] \mapsto \{u_1, \ldots, u_k\}{\sim}$.

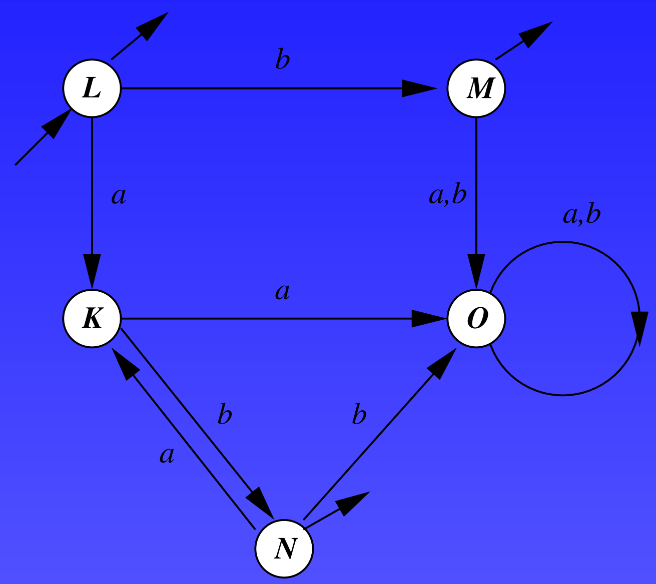
Proof We have to show that $\{u_1, \ldots, u_k\} \sim \{v_1, \ldots, v_l\}$ is equivalent to $[u_1, \ldots, u_k] = [v_1, \ldots, v_l]$.
Indeed, ...

Again, denote by $[u_1, \ldots, u_k]$ the transformation of U given by the set $\{u_1, \ldots, u_k\}$.

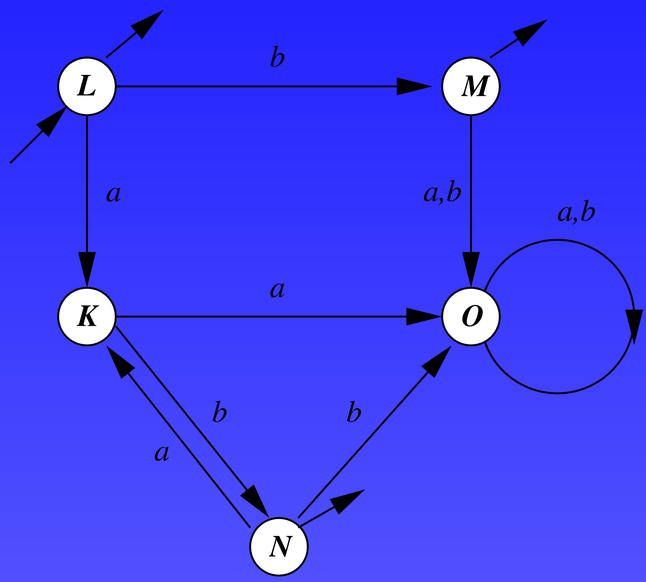
Theorem The transformation monoid $\{\ [u_1, \ldots, u_k] \mid k \geq 0,\ u_1, \ldots, u_k \in A^*\ \}$ with the operations of composition and $[u_1, \ldots, u_k] \vee [v_1, \ldots, v_l] = [u_1, \ldots, u_k, v_1, \ldots, v_l]$. is isomorphic to the syntactic semiring $(\mathsf{S}\,(L), \cdot, \vee)$ of the language L. An isomorphism is given by: $[u_1, \ldots, u_k] \mapsto \{u_1, \ldots, u_k\}\sim$.

Proof We have to show that $\{u_1, \ldots, u_k\} \sim \{v_1, \ldots, v_l\}$ is equivalent to $[u_1, \ldots, u_k] = [v_1, \ldots, v_l]$.
Indeed, ...

Example Let $A = \{a, b\}$ and $L = (ab)^* + b$. Denoting $K = a^{-1}L = b(ab)^*$, $M = b^{-1}L = 1$, $O = a^{-1}K = \emptyset$, $N = b^{-1}K = (ab)^*$, we see that the canonical minimal automaton $\mathcal{D}$ of L looks as follows.



see that the canonical minimal automaton $\mathcal{D}$ of L looks as follows.



Let us calculate the transformation matrix M of $\mathcal{D}$ first.

	L	K	M	N	O
1	L	K	M	N	O
a	K	O	O	K	O
b	M	N	O	O	O
a^2	O	O	O	O	O
ab	N	O	O	N	O
ba	O	K	O	O	O
bab	O	N	O	O	O

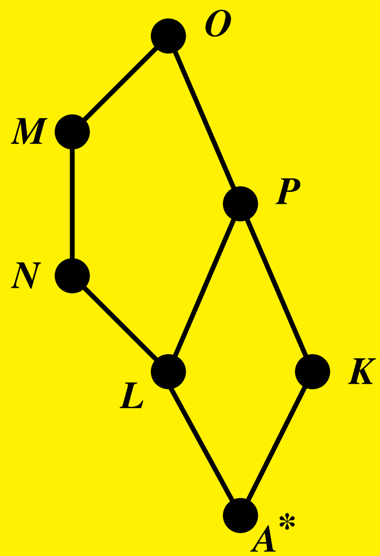
The syntactic monoid O has the presentation
$< \ a, b \ | \ a^2 = b^2 = 0, \ aba = a \ >$.

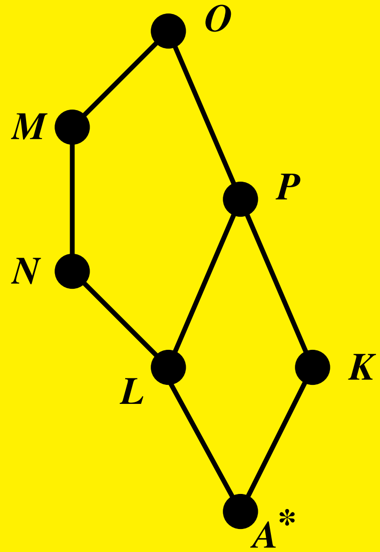
	L	K	M	N	O
1	L	K	M	N	O
a	K	O	O	K	O
b	M	N	O	O	O
a^2	O	O	O	O	O
ab	N	O	O	N	O
ba	O	K	O	O	O
bab	O	N	O	O	O

The syntactic monoid O has the presentation
$< \ a, b \ | \ a^2 = b^2 = 0, \ aba = a \ > \ .$
Now consider consecutively the matrices $\overline{M}$ - the **shadow** of M - we put 1 instead of terminal states and 0 otherwise, and the **right completion** rc$(\overline{M})$ of $\overline{M}$.

	L	K	M	N	O	A^*	$K \cap L$
1	1	0	1	1	0	1	0
a	0	0	0	0	0	1	0
b	1	1	0	0	0	1	1
a^2	0	0	0	0	0	1	0
ab	1	0	0	1	0	1	0
ba	0	0	0	0	0	1	0
bab	0	1	0	0	0	1	0

We see that U $= \{L, K, M, N, O, A^*, K \cap L\}$. Denoting $P = K \cap L$, the order on the states from U (i.e. the reverse inclusion) is

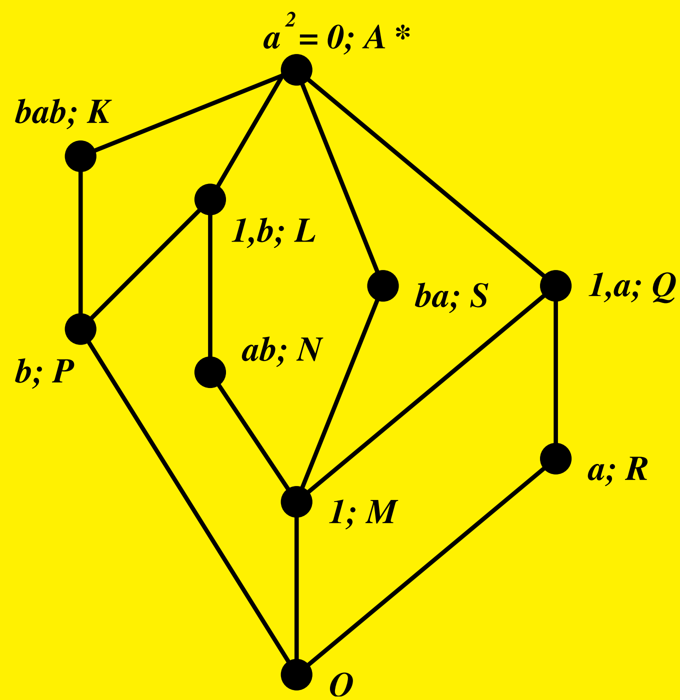

O
M
P
N
L
K
A*



We can proceed by calculating the matrices dc(M) (***down completion***) and rc(M) placing them into a single schema.

	L	K	M	N	O	A^*	P
1	L	K	M	N	O	A^*	P
a	K	O	O	K	O	A^*	O
b	M	N	O	O	O	A^*	M
a^2	O	O	O	O	O	A^*	O
ab	N	O	O	N	O	A^*	O
ba	O	K	O	O	O	A^*	O
bab	O	N	O	O	O	A^*	O
$\emptyset$	A^*	A^*	A^*	A^*	A^*		
$1, a$	P	O	O	O	O		
$1, b$	M	O	O	O	O		

The matrix dc(M) gives us the syntactic semiring S of L; its order reduct is depicted below. The parts of labels after semicolons correspond to the representation of the syntactic semiring by L-closed subsets. Here $Q = Lb^{-1} = (ab)^*a + 1,\ R = (ab)^{-1}Lb^{-1} = (ab)^*a,\ S = a^{-1}Lb^{-1} = b(ab)^*a + 1.$



Eilenberg-type theorems

Eilenberg 1976 : varieties of (regular) languages correspond to pseudovarieties of finite monoids; $L \mapsto (\mathsf{O}\,(L), \cdot)$.

Eilenberg-type theorems

Eilenberg 1976 : varieties of (regular) languages correspond to pseudovarieties of finite monoids; $L \mapsto (\mathsf{O}(L), \cdot)$.

Ex. star-free languages $\mapsto$ aperiodic monoids $=$ Mod $(x^{\omega+1} = x^{\omega})$

For a semigroup S and $a \in S$ there is exactly one idempotent in a subsemigroup of S generated by a; we denote it a^{ω}.

Eilenberg-type theorems

Eilenberg 1976 : varieties of (regular) languages correspond to pseudovarieties of finite monoids; $L \mapsto (\mathsf{O}\,(L), \cdot)$.
Ex. star-free languages $\mapsto$ aperiodic monoids $=$ Mod $(x^{\omega+1} = x^{\omega})$
For a semigroup S and $a \in S$ there is exactly one idempotent in a subsemigroup of S generated by a; we denote it a^{ω}.

Pin 1995 : positive varieties of languages correspond to pseudovarieties of finite ordered monoids; $L \mapsto (\mathsf{O}\,(L), \cdot, \leq)$.

Eilenberg-type theorems

Eilenberg 1976 : varieties of (regular) languages correspond to pseudovarieties of finite monoids; $L \mapsto (\mathsf{O}\,(L), \cdot)$.
Ex. star-free languages $\mapsto$ aperiodic monoids $=$ Mod $(x^{\omega+1} = x^{\omega})$
For a semigroup S and $a \in S$ there is exactly one idempotent in a subsemigroup of S generated by a; we denote it a^{ω}.

Pin 1995 : positive varieties of languages correspond to pseudovarieties of finite ordered monoids; $L \mapsto (\mathsf{O}\,(L), \cdot, \leq)$.
Ex. finite unions of $A^*a_1A^*a_2 \ldots a_kA^* \mapsto$ Mod $(x \leq 1)$

L.P. 1999 : conjunctive varieties of languages correspond to pseudovarieties of finite idempotent semirings; $L \mapsto (\mathsf{S}\,(L), \cdot, \vee)$.

Ex. complements of finite unions of $+$-languages $XA^* \cup A^*Y \cup Z$ where $X, Y, Z \subseteq A^+$ are finite $\mapsto$
Mod $(\ x^\omega y x^\omega = x^\omega, \ x^\omega y^\omega = x^\omega z \vee t y^\omega \ )$

Ex. complements of finite unions of $+$-languages $XA^* \cup A^*Y \cup Z$ where $X, Y, Z \subseteq A^+$ are finite $\mapsto$ Mod $(\ x^\omega y x^\omega = x^\omega,\ x^\omega y^\omega = x^\omega z \vee t y^\omega\ )$

Straubing 2002 : $\mathbb{C}$-varieties of languages correspond to $\mathbb{C}$-pseudovarieties of monoid homomorphisms;
$L \mapsto (\kappa_L : (A^*, \cdot) \twoheadrightarrow (\mathsf{O}\,(L), \cdot))$.

Ex. complements of finite unions of $+$-languages $XA^* \cup A^*Y \cup Z$ where $X, Y, Z \subseteq A^+$ are finite $\mapsto$
Mod ($x^\omega y x^\omega = x^\omega,\ x^\omega y^\omega = x^\omega z \vee t y^\omega$)

Straubing 2002 : $\mathbb{C}$-varieties of languages correspond to $\mathbb{C}$-pseudovarieties of monoid homomorphisms;
$L \mapsto (\kappa_L : (A^*, \cdot) \twoheadrightarrow (\mathsf{O}\,(L), \cdot))$.
Ex. certain classes of languages connected with circuit complexity classes $\mapsto$ all $\kappa : A^* \twoheadrightarrow O$ such that the stable subsemigroup of κ (the idempotent among $(\kappa(A))^n$ in the power monoid of O) belongs to a given pseudovariety of semigroups

Ex. complements of finite unions of $+$-languages $XA^* \cup A^*Y \cup Z$ where $X, Y, Z \subseteq A^+$ are finite $\mapsto$
Mod ($x^\omega y x^\omega = x^\omega,\ x^\omega y^\omega = x^\omega z \vee t y^\omega$)

Straubing 2002 : $\mathbb{C}$-varieties of languages correspond to $\mathbb{C}$-pseudovarieties of monoid homomorphisms;
$L \mapsto (\kappa_L : (A^*, \cdot) \twoheadrightarrow (\mathsf{O}\,(L), \cdot))$.
Ex. certain classes of languages connected with circuit complexity classes $\mapsto$ all $\kappa : A^* \twoheadrightarrow O$ such that the stable subsemigroup of κ (the idempotent among $(\kappa(A))^n$ in the power monoid of O) belongs to a given pseudovariety of semigroups

Positive $\mathbb{C}$-varieties of languages correspond to $\mathbb{C}$-pseudovarieties of ordered monoid homomorphisms;

$$L \mapsto (\kappa_L : (A^*, \cdot, =) \twoheadrightarrow (\mathsf{O}\,(L), \cdot, \leq)).$$

L.P. 2003 : $\mathbb{D}$-varieties of languages correspond to $\mathbb{D}$-pseudovarieties of idempotent semiring homomorphisms;

$$L \mapsto (\phi_L : (A^\square, \cdot, \cup) \twoheadrightarrow (\mathsf{S}\,(L), \cdot, \vee)).$$

$L \mapsto (\kappa_L : (A^*, \cdot, =) \twoheadrightarrow (\mathsf{O}\,(L), \cdot, \leq))$.

L.P. 2003 : $\mathbb{D}$-varieties of languages correspond to $\mathbb{D}$-pseudovarieties of idempotent semiring homomorphisms;
$L \mapsto (\phi_L : (A^{\square}, \cdot, \cup) \twoheadrightarrow (\mathsf{S}\,(L), \cdot, \vee))$.
It covers all previous results.
Ex. the so-called multiliteral varieties of languages.

$$L \mapsto (\kappa_L : (A^*, \cdot, =) \twoheadrightarrow (\mathsf{O}\,(L), \cdot, \leq)).$$

L.P. 2003 : $\mathbb{D}$-varieties of languages correspond to $\mathbb{D}$-pseudovarieties of idempotent semiring homomorphisms;
$L \mapsto (\phi_L : (A^\square, \cdot, \cup) \twoheadrightarrow (\mathsf{S}\,(L), \cdot, \vee)).$
It covers all previous results.
Ex. the so-called multiliteral varieties of languages.

$\mathbb{D}$-pseudovarieties of semiring homomorphisms

Let $\boldsymbol{S}$ be the class of all finite idempotent semirings. Let

$$\mathfrak{S} = \{\ \phi : A^\square \twoheadrightarrow S \mid A \text{ is a finite set and } S \in \boldsymbol{S}\ \}$$

be the class of all surjective semiring homomorphisms from a finitely generated free idempotent semirings onto a finite semirings.

We consider a category $\mathbb{D}$ of homomorphisms between finitely generated free idempotent semirings, i.e. objects are all semirings $A^\square$ where A is a finite set and the sets of morphisms $\mathbb{D}(B^\square, A^\square)$ consist of certain semiring homomorphisms from $B^\square$ to $A^\square$.

We consider a category $\mathbb{D}$ of homomorphisms between finitely generated free idempotent semirings, i.e. objects are all semirings $A^\square$ where A is a finite set and the sets of morphisms $\mathbb{D}(B^\square, A^\square)$ consist of certain semiring homomorphisms from $B^\square$ to $A^\square$.

Basic examples are the category $\mathbb{D}_{all}$ of all semiring homomorphisms, and the categories $\mathbb{D}_{mne}$, $\mathbb{D}_{mi}$, $\mathbb{D}_{ml}$ and $\mathbb{D}_l$ of all **multi-non-erasing**, **monoid induced**, **multi-literal** and **literal** homomorphisms, respectively. For finite alphabets A, B :
$f \in \mathbb{D}_{mne}(B^\square, A^\square)$ iff for each $b \in B$ there are $u_1, \ldots, u_k \in A^+$ such that $f(\{b\}) = \{u_1, \ldots, u_k\}$,
$f \in \mathbb{D}_{mi}(B^\square, A^\square)$ iff for each $b \in B$ there is $u \in A^*$ such that $f(\{b\}) = \{u\}$,
$f \in \mathbb{D}_{ml}(B^\square, A^\square)$ iff for each $b \in B$ there are $a_1, \ldots, a_k \in A$ such

that $f(\{b\}) = \{a_1, \ldots, a_k\}$, and
$f \in \mathbb{D}_l(B^\Box, A^\Box)$ iff for each $b \in B$ there is $a \in A$ such that
$f(\{b\}) = \{a\}$.

that $f(\{b\}) = \{a_1, \ldots, a_k\}$, and
$f \in \mathbb{D}_l(B^\square, A^\square)$ iff for each $b \in B$ there is $a \in A$ such that $f(\{b\}) = \{a\}$.

A class $\mathfrak{X} \subseteq \mathfrak{S}$ is a **$\mathbb{D}$-pseudovariety of semiring homomorphisms** if it satisfies :

(H) for each $(\phi : A^\square \twoheadrightarrow S) \in \mathfrak{X}$ and a surjective semiring homomorphism $\sigma : S \twoheadrightarrow T$ we have $\sigma\phi \in \mathfrak{X}$,

that $f(\{b\}) = \{a_1, \ldots, a_k\}$, and
$f \in \mathbb{D}_l(B^\square, A^\square)$ iff for each $b \in B$ there is $a \in A$ such that $f(\{b\}) = \{a\}$.

A class $\mathfrak{X} \subseteq \mathfrak{S}$ is a **$\mathbb{D}$-pseudovariety of semiring homomorphisms** if it satisfies :

(H) for each $(\phi : A^\square \twoheadrightarrow S) \in \mathfrak{X}$ and a surjective semiring homomorphism $\sigma : S \twoheadrightarrow T$ we have $\sigma\phi \in \mathfrak{X}$,

($S_{\mathbb{D}}$) for each $f \in \mathbb{D}(B^\square, A^\square)$ and $(\phi : A^\square \twoheadrightarrow S) \in \mathfrak{X}$ we have $(\ \phi f : B^\square \twoheadrightarrow \text{im}\,(\phi f)\ ) \in\ \mathfrak{X}$,

that $f(\{b\}) = \{a_1, \ldots, a_k\}$, and
$f \in \mathbb{D}_l(B^\Box, A^\Box)$ iff for each $b \in B$ there is $a \in A$ such that $f(\{b\}) = \{a\}$.

A class $\mathfrak{X} \subseteq \mathfrak{S}$ is a **$\mathbb{D}$-pseudovariety of semiring homomorphisms** if it satisfies :

(H) for each $(\phi : A^\Box \twoheadrightarrow S) \in \mathfrak{X}$ and a surjective semiring homomorphism $\sigma : S \twoheadrightarrow T$ we have $\sigma\phi \in \mathfrak{X}$,

($S_{\mathbb{D}}$) for each $f \in \mathbb{D}(B^\Box, A^\Box)$ and $(\phi : A^\Box \twoheadrightarrow S) \in \mathfrak{X}$ we have $(\ \phi f : B^\Box \twoheadrightarrow \mathrm{im}\,(\phi f)\ ) \in\ \mathfrak{X}$,

(P) for each non-negative integer m and for each system $\phi_1 : A^\Box \twoheadrightarrow S_1, \ldots, \phi_m : A^\Box \twoheadrightarrow S_m \in \mathfrak{X}$ we have $(\ (\phi_1, \ldots, \phi_m) : A^\Box \twoheadrightarrow \mathrm{im}\,(\phi_1, \ldots, \phi_m)\ ) \in \mathfrak{X}$

(here
$(\phi_1, \ldots, \phi_m)(U) = (\phi_1(U), \ldots, \phi_m(U)) \in S_1 \times \cdots \times S_m,\ U \in A^{\square}$,
and for $m = 0$ we have $(A^{\square} \twoheadrightarrow \{1\}) \in \mathfrak{X}$).

(here
$(\phi_1, \ldots, \phi_m)(U) = (\phi_1(U), \ldots, \phi_m(U)) \in S_1 \times \cdots \times S_m,\ U \in A^{\square}$,
and for $m = 0$ we have $(A^{\square} \twoheadrightarrow \{1\}) \in \mathfrak{X}$).

For finite sets A and B, a semiring homomorphism
$f : (B^{\square}, \cdot, \cup) \to (A^{\square}, \cdot, \cup)$ and $K \subseteq A^*,\ L \subseteq B^*$ we define

$$f^{[-1]}(K) = \{\ v \in B^* \mid f(\{v\}) \subseteq K\ \}\ ,$$

$$f^{(-1)}(K) = \{\ v \in B^* \mid f(\{v\}) \cap K \neq \emptyset\ \}\ ,$$
$$f(L) = \bigcup \{\ f(\{v\}) \mid v \in L\ \}\ .$$

Further, for the complement K^{c} of a language $K \subseteq A^*$, it holds

$$f^{[-1]}(K^{\mathsf{c}}) = (f^{(-1)}(K))^{\mathsf{c}}\ .$$

For a multiliteral homomorphism $f : B^\Box \longrightarrow A^\Box$ we define the **dual** multiliteral homomorphism $\widehat{f} : A^\Box \longrightarrow B^\Box$ by $b \in \widehat{f}(\{a\})$ iff $a \in f(\{b\})$ for $a \in A,\ b \in B$. Notice that $f^{(-1)}(K) = \widehat{f}(K)$ for each $K \subseteq A^*$.

For a multiliteral homomorphism $f : B^{\Box} \longrightarrow A^{\Box}$ we define the **dual** multiliteral homomorphism $\widehat{f} : A^{\Box} \rightarrow B^{\Box}$ by $b \in \widehat{f}(\{a\})$ iff $a \in f(\{b\})$ for $a \in A,\ b \in B$. Notice that $f^{(-1)}(K) = \widehat{f}(K)$ for each $K \subseteq A^*$.

A **class** of (regular) languages is an operator $\mathcal{L}$ assigning to every finite set A a set $\mathcal{L}(A)$ of regular languages over the alphabet A.

For a multiliteral homomorphism $f : B^{\Box} \to A^{\Box}$ we define the **dual** multiliteral homomorphism $\widehat{f} : A^{\Box} \to B^{\Box}$ by $b \in \widehat{f}(\{a\})$ iff $a \in f(\{b\})$ for $a \in A,\ b \in B$. Notice that $f^{(-1)}(K) = \widehat{f}(K)$ for each $K \subseteq A^*$.

A **class** of (regular) languages is an operator $\mathcal{L}$ assigning to every finite set A a set $\mathcal{L}(A)$ of regular languages over the alphabet A. Such a class is a **conjunctive variety** if
(i) each $\mathcal{L}(A)$ is closed with respect to finite intersections and quotients, and
(ii) for each finite sets A and B and $f : B^{\Box} \to A^{\Box},\ K \in \mathcal{L}(A)$ implies $f^{[-1]}(K) \in \mathcal{L}(B)$.

For a multiliteral homomorphism $f : B^\square \to A^\square$ we define the **dual** multiliteral homomorphism $\widehat{f} : A^\square \to B^\square$ by $b \in \widehat{f}(\{a\})$ iff $a \in f(\{b\})$ for $a \in A,\ b \in B$. Notice that $f^{(-1)}(K) = \widehat{f}(K)$ for each $K \subseteq A^*$.

A **class** of (regular) languages is an operator $\mathcal{L}$ assigning to every finite set A a set $\mathcal{L}(A)$ of regular languages over the alphabet A. Such a class is a **conjunctive variety** if
(i) each $\mathcal{L}(A)$ is closed with respect to finite intersections and quotients, and
(ii) for each finite sets A and B and $f : B^\square \to A^\square,\ K \in \mathcal{L}(A)$ implies $f^{[-1]}(K) \in \mathcal{L}(B)$.

It is a **$\mathbb{D}$-variety** if (i) is true and (ii) is satisfied for $f \in \mathbb{D}(B^\square, A^\square)$.

Finally, a class $\mathcal{L}$ is a **multiliteral variety** if
(i') each $\mathcal{L}(A)$ is closed with respect to finite unions and quotients, and
(ii') for each finite sets A and B and $f \in \mathbb{D}_{ml}(A^{\Box}, B^{\Box})$, $K \in \mathcal{L}(A)$ implies $f(K) \in \mathcal{L}(B)$.

Finally, a class $\mathcal{L}$ is a **multiliteral variety** if
(i') each $\mathcal{L}(A)$ is closed with respect to finite unions and quotients, and
(ii') for each finite sets A and B and $f \in \mathbb{D}_{ml}(A^{\Box}, B^{\Box})$, $K \in \mathcal{L}(A)$ implies $f(K) \in \mathcal{L}(B)$.

Define $\mathcal{L}^{\mathsf{c}}$ by $\mathcal{L}^{\mathsf{c}}(A) = \{\ L^{\mathsf{c}} \mid L \in \mathcal{L}(A)\ \}$.

Finally, a class $\mathcal{L}$ is a **multiliteral variety** if
(i') each $\mathcal{L}(A)$ is closed with respect to finite unions and quotients, and
(ii') for each finite sets A and B and $f \in \mathbb{D}_{ml}(A^{\square}, B^{\square})$, $K \in \mathcal{L}(A)$ implies $f(K) \in \mathcal{L}(B)$.

Define $\mathcal{L}^{\mathsf{c}}$ by $\mathcal{L}^{\mathsf{c}}(A) = \{\ L^{\mathsf{c}} \mid L \in \mathcal{L}(A)\ \}$.

Fix a category $\mathbb{D}$. We can assign to any class of languages $\mathcal{L}$ the pseudovariety

$$\mathsf{S}\,(\mathcal{L}) = \langle\ \{\ (\mathsf{S}\,(L), \cdot, \vee) \mid A \text{ a finite set, } L \in \mathcal{L}(A)\ \}\ \rangle_{\mathcal{S}}$$

of idempotent semirings generated by all syntactic semirings of members of $\mathcal{L}$,

members of $\mathcal{L}$, and the $\mathbb{D}$-pseudovariety

$$\mathbf{S}_{\mathbb{D}}(\mathcal{L}) = \langle \; \{ \; \phi_L \mid A \text{ a finite set, } L \in \mathcal{L}(A) \; \} \; \rangle_{\mathfrak{S},\mathbb{D}}$$

of idempotent semiring homomorphisms generated by all syntactic semiring homomorphisms of members of $\mathcal{L}$.

members of $\mathcal{L}$, and the $\mathbb{D}$-pseudovariety

$$\mathbf{S}_{\mathbb{D}}(\mathcal{L}) = \langle\ \{\ \phi_L \mid A \text{ a finite set, } L \in \mathcal{L}(A)\ \}\ \rangle_{\mathfrak{S},\mathbb{D}}$$

of idempotent semiring homomorphisms generated by all syntactic semiring homomorphisms of members of $\mathcal{L}$.

Conversely, for a class $\mathcal{X}$ of idempotent semirings and a finite set A, we put

$$(\mathsf{L}\,(\mathcal{X}))(A) = \{\ L \subseteq A^* \mid (\mathsf{S}\,(L), \cdot, \vee) \in \mathcal{X}\ \}$$

members of $\mathcal{L}$, and the $\mathbb{D}$-pseudovariety

$$\mathbf{S}_{\,\mathbb{D}}(\mathcal{L}) = \langle\ \{\ \phi_L \mid A \text{ a finite set, } L \in \mathcal{L}(A)\ \}\ \rangle_{\mathfrak{S},\mathbb{D}}$$

of idempotent semiring homomorphisms generated by all syntactic semiring homomorphisms of members of $\mathcal{L}$.

Conversely, for a class $\mathcal{X}$ of idempotent semirings and a finite set A, we put

$$(\mathsf{L}\,(\mathcal{X}))(A) = \{\ L \subseteq A^* \mid (\mathsf{S}\,(L), \cdot, \vee) \in \mathcal{X}\ \}$$

and similarly for a class $\mathfrak{X}$ of idempotent semiring homomorphisms and a finite set A, we put

$$(\mathbf{L}\,(\mathfrak{X}))(A) = \{\, L \subseteq A^* \mid \phi_L \in \mathfrak{X}\,\},\ (\mathbf{L}^{\mathsf{c}}(\mathfrak{X}))(A) = \{\, L \subseteq A^* \mid \phi_{L^{\mathsf{c}}} \in \mathfrak{X}\,\}$$

Theorem (i) The assignments $\mathcal{L} \mapsto \mathsf{S}(\mathcal{L})$ and $\mathcal{X} \mapsto \mathsf{L}(\mathcal{X})$ are mutually inverse bijections between the conjunctive varieties of languages and pseudovarieties of finite idempotent semirings.

Theorem (i) The assignments $\mathcal{L} \mapsto \mathsf{S}(\mathcal{L})$ and $\mathcal{X} \mapsto \mathsf{L}(\mathcal{X})$ are mutually inverse bijections between the conjunctive varieties of languages and pseudovarieties of finite idempotent semirings.
(ii) The assignments $\mathcal{L} \mapsto \mathbf{S}_{\mathbb{D}}(\mathcal{L})$ and $\mathfrak{X} \mapsto \mathbf{L}(\mathfrak{X})$ are mutually inverse bijections between the $\mathbb{D}$-varieties of languages and $\mathbb{D}$-pseudovarieties of homomorphisms of idempotent semirings.

Theorem (i) The assignments $\mathcal{L} \mapsto \mathsf{S}(\mathcal{L})$ and $\mathcal{X} \mapsto \mathsf{L}(\mathcal{X})$ are mutually inverse bijections between the conjunctive varieties of languages and pseudovarieties of finite idempotent semirings.
(ii) The assignments $\mathcal{L} \mapsto \mathbf{S}_{\mathbb{D}}(\mathcal{L})$ and $\mathfrak{X} \mapsto \mathbf{L}(\mathfrak{X})$ are mutually inverse bijections between the $\mathbb{D}$-varieties of languages and $\mathbb{D}$-pseudovarieties of homomorphisms of idempotent semirings.
(iii) The assignments $\mathcal{L} \mapsto \mathbf{S}_{\mathbb{D}_{ml}}(\mathcal{L}^{\mathsf{c}})$ and $\mathfrak{X} \mapsto \mathbf{L}^{\mathsf{c}}(\mathfrak{X})$ are mutually inverse bijections between the multiliteral varieties of languages and $\mathbb{D}_{ml}$-pseudovarieties of homomorphisms of idempotent semirings.

Theorem (i) The assignments $\mathcal{L} \mapsto \mathsf{S}\,(\mathcal{L})$ and $\mathcal{X} \mapsto \mathsf{L}\,(\mathcal{X})$ are mutually inverse bijections between the conjunctive varieties of languages and pseudovarieties of finite idempotent semirings.
(ii) The assignments $\mathcal{L} \mapsto \mathbf{S}_{\mathbb{D}}(\mathcal{L})$ and $\mathfrak{X} \mapsto \mathbf{L}\,(\mathfrak{X})$ are mutually inverse bijections between the $\mathbb{D}$-varieties of languages and $\mathbb{D}$-pseudovarieties of homomorphisms of idempotent semirings.
(iii) The assignments $\mathcal{L} \mapsto \mathbf{S}_{\mathbb{D}_{ml}}(\mathcal{L}^{\mathsf{c}})$ and $\mathfrak{X} \mapsto \mathbf{L}^{\mathsf{c}}(\mathfrak{X})$ are mutually inverse bijections between the multiliteral varieties of languages and $\mathbb{D}_{ml}$-pseudovarieties of homomorphisms of idempotent semirings.

Examples of multiliteral varieties of languages

We define the classes of languages $\mathcal{S}(n)$, $\mathcal{H}_{1/2}(k)$, $\mathcal{H}_{3/2}(k)$, $\mathcal{H}_{1/2}$, $\mathcal{H}_{3/2}$, $\mathcal{L}(k,m,l)$, $\mathcal{L}(m)$ below. It is obvious that all of them

are multiliteral varieties of languages.

For a finite set A and $n, k, l \geq 0,\ m \geq 1$ put
$\mathcal{S}(n)(A) =$ class of all languages over A of star height $\leq n$,

are multiliteral varieties of languages.

For a finite set A and $n, k, l \geq 0,\ m \geq 1$ put
$\mathcal{S}(n)(A) =$ class of all languages over A of star height $\leq n$,
$\mathcal{H}_{1/2}(k)(A) =$ all finite unions of $A^*a_1A^*a_2\ldots a_pA^*$, where
$0 \leq p \leq k,\ a_1, \ldots, a_p \in A$,

are multiliteral varieties of languages.

For a finite set A and $n, k, l \geq 0,\ m \geq 1$ put
$\mathcal{S}(n)(A) =$ class of all languages over A of star height $\leq n$,
$\mathcal{H}_{1/2}(k)(A) =$ all finite unions of $A^*a_1A^*a_2\dots a_pA^*$, where
$0 \leq p \leq k,\ a_1, \dots, a_p \in A$,
$\mathcal{H}_{3/2}(k)(A) =$ all finite unions of $B_0^*a_1B_1^*a_2\dots a_pB_p^*$, where
$0 \leq p \leq k,\ a_1, \dots, a_p \in A$ and $B_0, \dots, B_p \subseteq A$,

are multiliteral varieties of languages.

For a finite set A and $n, k, l \geq 0,\ m \geq 1$ put
$\mathcal{S}(n)(A) =$ class of all languages over A of star height $\leq n$,
$\mathcal{H}_{1/2}(k)(A) =$ all finite unions of $A^*a_1A^*a_2\dots a_pA^*$, where
$0 \leq p \leq k,\ a_1, \dots, a_p \in A$,
$\mathcal{H}_{3/2}(k)(A) =$ all finite unions of $B_0^*a_1B_1^*a_2\dots a_pB_p^*$, where
$0 \leq p \leq k,\ a_1, \dots, a_p \in A$ and $B_0, \dots, B_p \subseteq A$,
$\mathcal{L}(k, m, l) =$ all finite unions of $uB_1^*\dots B_m^*v$, where
$u, v \in A^*, |u| \leq k, |v| \leq\ l,\ B_1, \dots, B_m \subseteq A$.

are multiliteral varieties of languages.

For a finite set A and $n, k, l \geq 0,\ m \geq 1$ put
$\mathcal{S}(n)(A) =$ class of all languages over A of star height $\leq n$,
$\mathcal{H}_{1/2}(k)(A) =$ all finite unions of $A^*a_1A^*a_2 \ldots a_pA^*$, where
$0 \leq p \leq k,\ a_1, \ldots, a_p \in A$,
$\mathcal{H}_{3/2}(k)(A) =$ all finite unions of $B_0^*a_1B_1^*a_2 \ldots a_pB_p^*$, where
$0 \leq p \leq k,\ a_1, \ldots, a_p \in A$ and $B_0, \ldots, B_p \subseteq A$,
$\mathcal{L}(k, m, l) =$ all finite unions of $uB_1^* \ldots B_m^*v$, where
$u, v \in A^*, |u| \leq k, |v| \leq\ l,\ B_1, \ldots, B_m \subseteq A$.
Further, put $\mathcal{H}_{1/2}(A) = \bigcup_{k \geq 0} \mathcal{H}_{1/2}(k)(A)$,

$$\mathcal{H}_{3/2}(A) = \bigcup_{k \geq 0} \mathcal{H}_{3/2}(k)(A),\ \mathcal{L}(m)(A) = \bigcup_{k,l \geq 0} \mathcal{L}(k, m, l)(A)\ .$$

The varieties $\mathcal{H}_{1/2}$ and $\mathcal{H}_{3/2}$ are members of the so-called Straubing-Thérien hierarchy.

Pseudoidentities

This section modifies the approach by Kunc concerning pseudovarieties of monoid homomorphisms to the case of semirings.

Pseudoidentities

This section modifies the approach by Kunc concerning pseudovarieties of monoid homomorphisms to the case of semirings.

Recall that $(\pi_S)_{S \in \boldsymbol{S}}$ is an ***n*-ary implicit operation** $(n \geq 0)$ for the class $\boldsymbol{S}$ of all finite idempotent semirings if $\pi_S : S^n \to S$ $(S \in \boldsymbol{S})$ is a mapping and for each semiring homomorphism $\sigma : S \to T$ and $s_1, \ldots, s_n \in S$ we have

$$\sigma(\pi_S(s_1, \ldots, s_n)) = \pi_T(\sigma(s_1), \ldots, \sigma(s_n)) \ .$$

Pseudoidentities

This section modifies the approach by Kunc concerning pseudovarieties of monoid homomorphisms to the case of semirings.

Recall that $(\pi_S)_{S\in\boldsymbol{S}}$ is an n-**ary implicit operation** $(n \geq 0)$ for the class $\boldsymbol{S}$ of all finite idempotent semirings if $\pi_S : S^n \to S$ $(S \in \boldsymbol{S})$ is a mapping and for each semiring homomorphism $\sigma : S \to T$ and $s_1, \ldots, s_n \in S$ we have

$$\sigma(\pi_S(s_1, \ldots, s_n)) = \pi_T(\sigma(s_1), \ldots, \sigma(s_n)) \ .$$

An n-**ary pseudoidentity** is an ordered pair $\pi = \rho$ of n-ary implicit operations. Let $v_1, v_2, \ldots$ be a fixed sequence of pairwise different variables and let $V_n = \{v_1, \ldots, v_n\}$ for each $n \geq 0$.

Let $\mathbb{D}$ be a category of of homomorphisms of free finitely generated idempotent semirings. The pseudoidentity $\pi = \rho$ is **$\mathbb{D}$-satisfied** in a semiring homomorphism $(\phi : A^{\square} \twoheadrightarrow S) \in \mathfrak{S}$ if for each $f \in \mathbb{D}(V_n^{\square}, A^{\square})$ we have

$$\pi_S((\phi f)(v_1), \ldots, (\phi f)(v_n)) = \rho_S((\phi f)(v_1), \ldots, (\phi f)(v_n)) \; .$$

Let $\mathbb{D}$ be a category of of homomorphisms of free finitely generated idempotent semirings. The pseudoidentity $\pi = \rho$ is **$\mathbb{D}$-satisfied** in a semiring homomorphism $(\phi : A^{\square} \twoheadrightarrow S) \in \mathfrak{S}$ if for each $f \in \mathbb{D}(V_n^{\square}, A^{\square})$ we have

$$\pi_S((\phi f)(v_1), \ldots, (\phi f)(v_n)) = \rho_S((\phi f)(v_1), \ldots, (\phi f)(v_n)) \ .$$

Reiterman-type theorem

Theorem Let $\mathbb{D}$ be a category of homomorphisms of finitely generated free idempotent semirings containing all bijections. Then a class $\mathfrak{X} \subseteq \mathfrak{S}$ is a $\mathbb{D}$-pseudovariety if and only if there exists a set Σ of pseudoidentities such that $\mathfrak{X} = \mathrm{Mod}_{\mathbb{D}}(\Sigma)$.

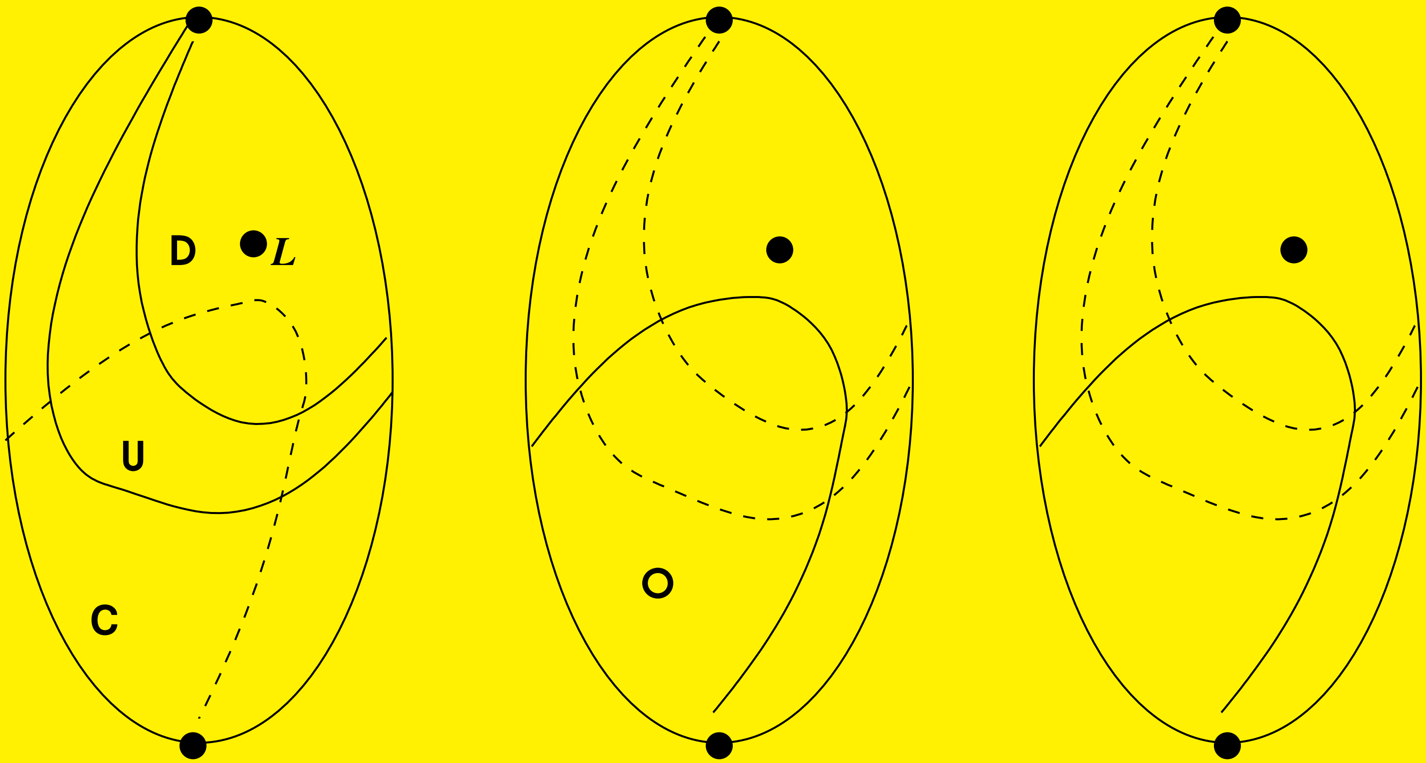
Three models of syntactic semirings

We put : C $=$
$\{\ u_1^{-1}Lv_1^{-1} \cap \cdots \cap u_k^{-1}Lv_k^{-1} \mid k \geq 0,\ u_1, \ldots, u_k, v_1, \ldots, v_k \in A^*\ \}$
(we get A^* for $k = 0$).

A subset X of A^* is called ***L*-closed** if
$(\ \forall\ u_1, \ldots, u_k \in X\ )\ (\ \forall\ v \in A^*\ )$
$(\ \{u_1, \ldots, u_k, v\} \sim \{u_1, \ldots, u_k\}$ implies $v \in X\ )$.

Proposition The elements of C are exactly the L-closed sets and the closure of a set $X \subseteq A^*$ is
$X^L = \bigcap \{\ x^{-1}Ly^{-1} \mid X \subseteq x^{-1}Ly^{-1},\ x, y \in A^*\ \}$.

$(\mathsf{C}, \subseteq) \qquad \cong \qquad (\mathsf{S}, \leq) \qquad \cong \qquad (\mathsf{T}, \leq)$



$\emptyset^L \qquad\qquad \emptyset\sim \qquad\qquad$ the constant map on A^*

$$Q \quad \overset{\chi}{\longmapsto} \quad \bigvee \{\ \{u\}\sim \mid u \in Q\ \}$$

$$\{u_1, \ldots, u_k\}^L \quad \overset{\psi}{\longleftarrow} \quad \{u_1, \ldots, u_k\}\sim \quad \longleftrightarrow \quad [u_1, \ldots, u_k]$$

Language equations

Let $\mathcal{L} = (2^{A^*}, \emptyset, \{a_1\}, \ldots, \{a_n\}, \cdot, \cup, ^*)$. For a given $r \in \text{Reg}\,(A, X)$ and a regular language L over A we will consider the inequality

$$r(x_1, \ldots, x_m) \subseteq L \tag{$*$}$$

and the equation

$$r(x_1, \ldots, x_m) = L \;. \tag{$**$}$$

Language equations

Let $\mathcal{L} = (2^{A^*}, \emptyset, \{a_1\}, \ldots, \{a_n\}, \cdot, \cup, ^*)$. For a given $r \in \mathsf{Reg}\,(A, X)$ and a regular language L over A we will consider the inequality

$$r(x_1, \ldots, x_m) \subseteq L \qquad (*)$$

and the equation

$$r(x_1, \ldots, x_m) = L\ . \qquad (**)$$

An m-tuple $(P_1, \ldots, P_m)$ of languages over A is called a **solution** of the inequality $(*)$ in $\mathcal{L}$ if $\overline{r}(P_1, \ldots, P_m) \subseteq L$. Similarly, it is a solution of $(**)$ in $\mathcal{L}$ if $\overline{r}(P_1, \ldots, P_m) = L$.

A solution $(P_1, \ldots, P_m)$ of $(*)$ in $\mathcal{L}$ is **maximal** if for every other solution $(Q_1, \ldots, Q_m)$,

$$P_1 \subseteq Q_1, \ldots, P_m \subseteq Q_m \text{ implies } P_1 = Q_1, \ldots, P_m = Q_m \; .$$

Similarly for the equation $(**)$.

A solution $(P_1, \ldots, P_m)$ of $(*)$ in $\mathcal{L}$ is **maximal** if for every other solution $(Q_1, \ldots, Q_m)$,

$$P_1 \subseteq Q_1, \ldots, P_m \subseteq Q_m \text{ implies } P_1 = Q_1, \ldots, P_m = Q_m \ .$$

Similarly for the equation $(**)$.

Theorem (a reformulation of Conway's Theorem VI.9.)
Let $(P_1, \ldots, P_m)$ be a solution of the inequality $r(x_1, \ldots, x_m) \subseteq L$ in $\mathcal{L}$. Then $(P_1^L, \ldots, P_m^L)$ is again a solution of this inequality. Consequently, the components of any maximal solution are L-closed; in particular, they are regular languages.

A solution $(P_1, \ldots, P_m)$ of $(*)$ in $\mathcal{L}$ is **maximal** if for every other solution $(Q_1, \ldots, Q_m)$,

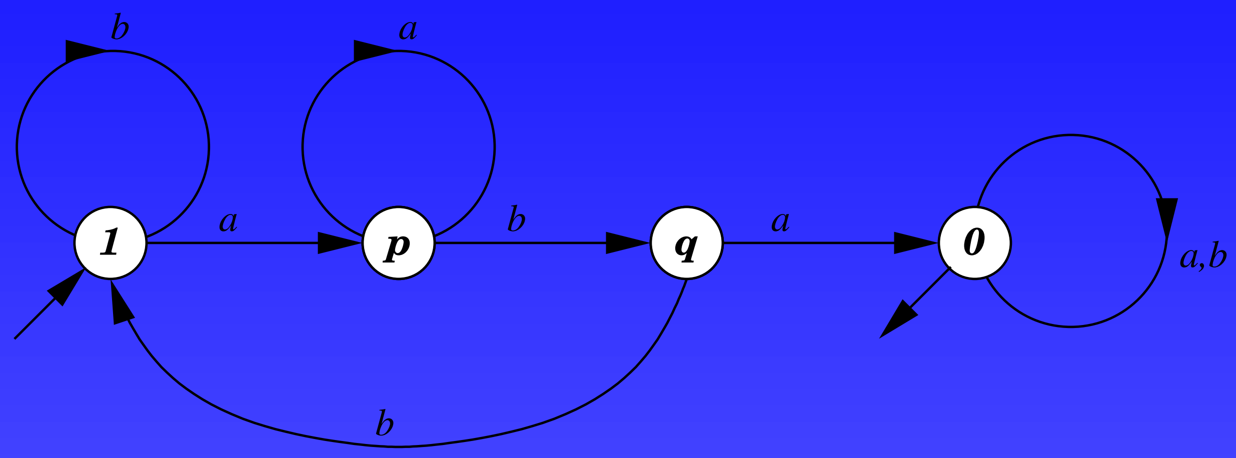
$$P_1 \subseteq Q_1, \ldots, P_m \subseteq Q_m \text{ implies } P_1 = Q_1, \ldots, P_m = Q_m \ .$$

Similarly for the equation $(**)$.

Theorem (a reformulation of Conway's Theorem VI.9.)
Let $(P_1, \ldots, P_m)$ be a solution of the inequality $r(x_1, \ldots, x_m) \subseteq L$ in $\mathcal{L}$. Then $(P_1^L, \ldots, P_m^L)$ is again a solution of this inequality. Consequently, the components of any maximal solution are L-closed; in particular, they are regular languages.

Example
Let $A = \{a, b\}$ and consider the language $L = A^*abaA^*$.

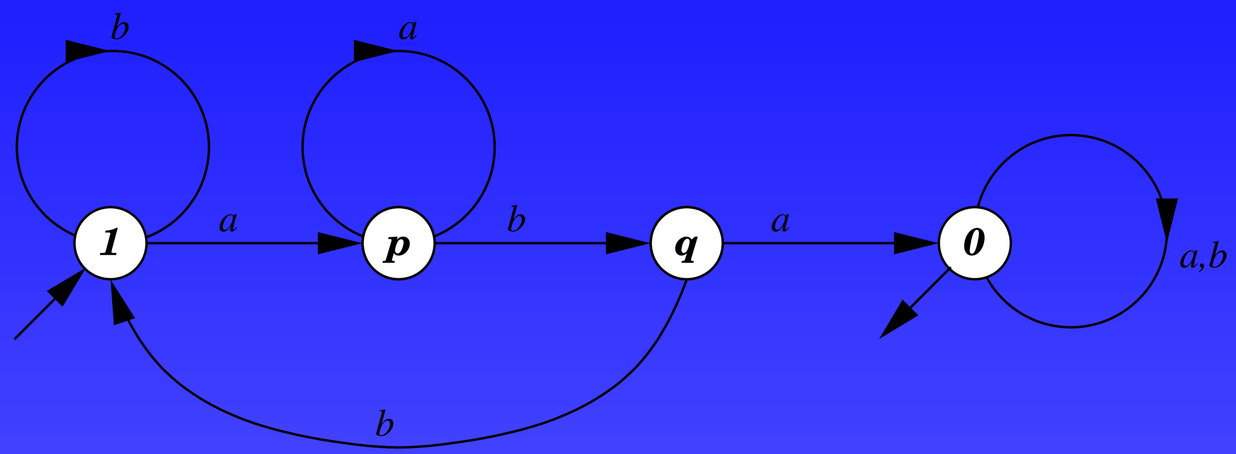
Denoting $\boldsymbol{1} = L$, $\boldsymbol{p} = L \cup baA^*$, $\boldsymbol{q} = L \cup aA^*$, $\boldsymbol{0} = A^*$, we see that the minimal automaton of L looks as follows.



The order on states is given by

$\boldsymbol{0} < \boldsymbol{p} < \boldsymbol{1}$, $\boldsymbol{0} < \boldsymbol{q} < \boldsymbol{1}$, $\boldsymbol{p}, \boldsymbol{q}$ incomparable. Note that $\boldsymbol{p} \cap \boldsymbol{q} = \boldsymbol{1}$.

Denoting $\boldsymbol{1} = L,\ \boldsymbol{p} = L \cup baA^*,\ \boldsymbol{q} = L \cup aA^*,\ \boldsymbol{0} = A^*$, we see that the minimal automaton of L looks as follows.



The order on states is given by
$\boldsymbol{0} < \boldsymbol{p} < \boldsymbol{1},\ \boldsymbol{0} < \boldsymbol{q} < \boldsymbol{1},\ \boldsymbol{p}, \boldsymbol{q}$ incomparable. Note that $\boldsymbol{p} \cap \boldsymbol{q} = \boldsymbol{1}$.

Now we calculate the transitions given by $1, a, b, a^2, ab, ba, b^2, a^3, \ldots$ and regular expressions for corresponding $\approx$-classes.

u	$\boldsymbol{1}$ $\boldsymbol{p}$ $\boldsymbol{q}$ $\boldsymbol{0}$	$u \approx$
1	$\boldsymbol{1}$ $\boldsymbol{p}$ $\boldsymbol{q}$ $\boldsymbol{0}$	1
a	$\boldsymbol{p}$ $\boldsymbol{p}$ $\boldsymbol{0}$ $\boldsymbol{0}$	$a^+b^2(b + a^+b^2)^* + a^+$
b	$\boldsymbol{1}$ $\boldsymbol{q}$ $\boldsymbol{1}$ $\boldsymbol{0}$	b
ab	$\boldsymbol{q}$ $\boldsymbol{q}$ $\boldsymbol{0}$ $\boldsymbol{0}$	$a^+b^2(b + a^+b^2)^*a^+b + a^+b$
ba	$\boldsymbol{p}$ $\boldsymbol{0}$ $\boldsymbol{p}$ $\boldsymbol{0}$	$b(b + a^+b^2)^*a^+$
b^2	$\boldsymbol{1}$ $\boldsymbol{1}$ $\boldsymbol{1}$ $\boldsymbol{0}$	$b^2(b + a^+b^2)^*$
aba	$\boldsymbol{0}$ $\boldsymbol{0}$ $\boldsymbol{0}$ $\boldsymbol{0}$	A^*abaA^*
ab^2	$\boldsymbol{1}$ $\boldsymbol{1}$ $\boldsymbol{0}$ $\boldsymbol{0}$	$a^+b^2(b + a^+b^2)^*$
bab	$\boldsymbol{q}$ $\boldsymbol{0}$ $\boldsymbol{q}$ $\boldsymbol{0}$	$ba^+b^2(b + a^+b^2)^*a^+b + ba^+b$
b^2a	$\boldsymbol{p}$ $\boldsymbol{p}$ $\boldsymbol{p}$ $\boldsymbol{0}$	$b^2(b + a^+b^2)^*a^+$
bab^2	$\boldsymbol{1}$ $\boldsymbol{0}$ $\boldsymbol{1}$ $\boldsymbol{0}$	$ba^+b^2(b + a^+b^2)^*$
b^2ab	$\boldsymbol{q}$ $\boldsymbol{q}$ $\boldsymbol{q}$ $\boldsymbol{0}$	$b^2(b + a^+b^2)^*a^+b$

The syntactic monoid of L has the presentation

$$< \ a, b \ \mid a^2 = a, \ b^3 = b^2, \ aba = 0, \ ab^2a = a, \ b^2ab^2 = b^2 \ >$$

and its order is given componentwise according to the second column.

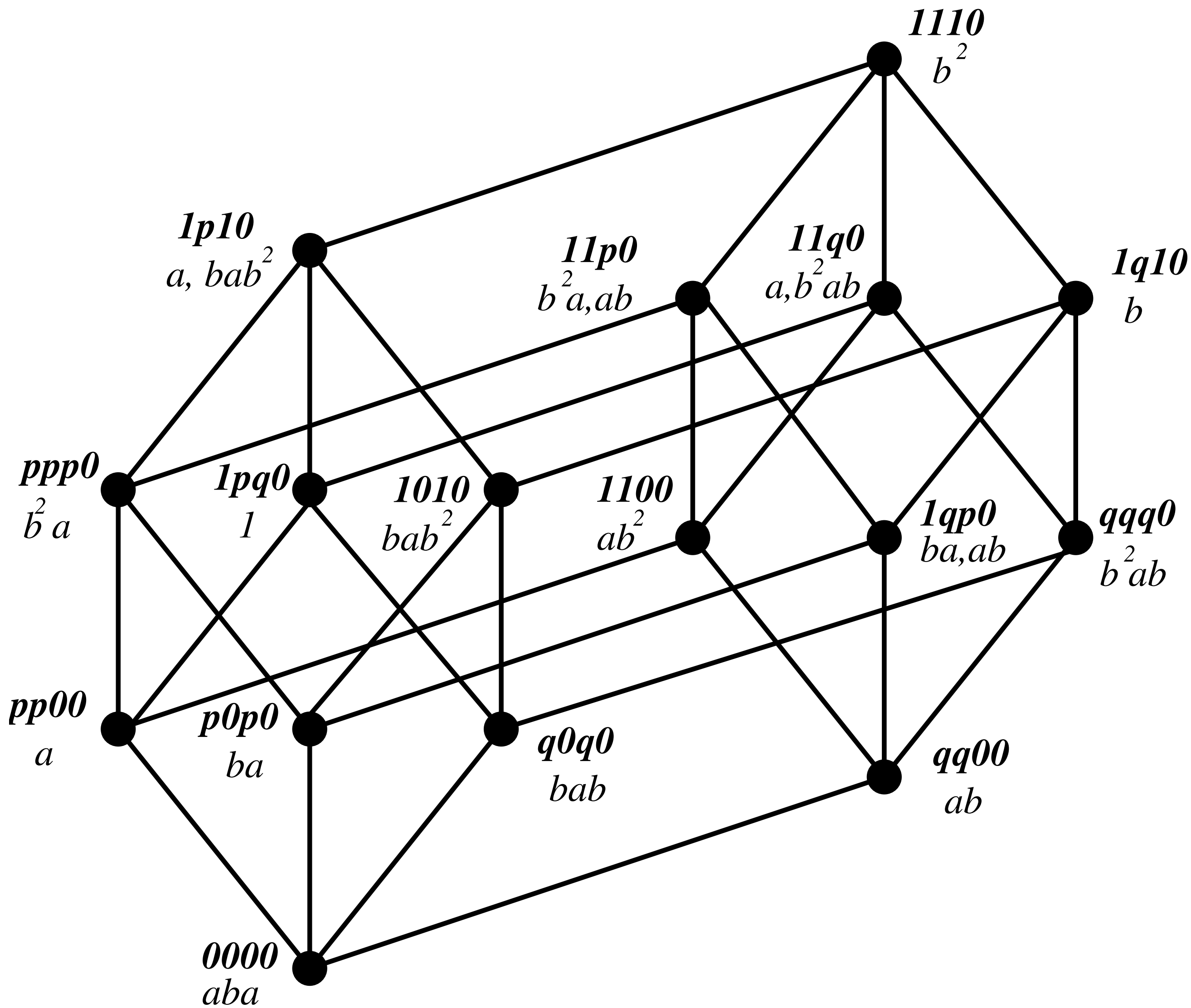
The syntactic monoid of L has the presentation

$$< \ a, b \ \mid a^2 = a, \ b^3 = b^2, \ aba = 0, \ ab^2a = a, \ b^2ab^2 = b^2 \ >$$

and its order is given componentwise according to the second column.

Some hereditary subsets of $(\mathsf{O}\,, \leq)$ act on U in the same way as some elements of O, for instance $[a, ba, aba] = [b^2a]$, some subsets give rise to new transformations:

$[a, b^2ab] : (\boldsymbol{1}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{0}) \mapsto (\boldsymbol{1}, \boldsymbol{1}, \boldsymbol{q}, \boldsymbol{0})$. This gives the label $\begin{matrix} \boldsymbol{1}\ \boldsymbol{1}\ \boldsymbol{q}\ \boldsymbol{0} \\ a,\ b^2ab \end{matrix}$ in the order reduct of the syntactic semiring of L depicted below.


1110
b2
1p10
a, bab2
11p0
b2a,ab
11q0
a,b2ab
1q10
b
ppp0
b2a
1pq0
1
1010
bab2
1100
ab2
1qp0
ba,ab
qqq0
b2ab
pp00
a
p0p0
ba
q0q0
bab
qq00
ab
0000
aba

Consider the inequality $x^2 \leq L$ and the equation $x^2 = L$. Write $\{u_1, \ldots, u_k\}$ instead of $\{u_1, \ldots, u_k\}\sim$. The elements of S with $x^2 \leq \chi(L)$ are exactly $\{aba\}$, $\{ab\}$, $\{ba\}$ (not, for instance, $\{ab, ba\}$ since $\{ab, ba\}^2 = \{bab, a, aba\} = \{1\}$). The last two are maximal solutions.

Consider the inequality $x^2 \leq L$ and the equation $x^2 = L$. Write $\{u_1, \ldots, u_k\}$ instead of $\{u_1, \ldots, u_k\}\sim$. The elements of S with $x^2 \leq \chi(L)$ are exactly $\{aba\}$, $\{ab\}$, $\{ba\}$ (not, for instance, $\{ab, ba\}$ since $\{ab, ba\}^2 = \{bab, a, aba\} = \{1\}$). The last two are maximal solutions.

The corresponding languages are
$\psi(\{ab\}) = ab \approx \cup\ aba \approx = a^+b^2(b + a^+b^2)^*a^+b + a^+b + A^*abaA^*$,
$\psi(\{ba\}) = ba \approx \cup\ aba \approx = b(b + a^+b^2)^*a^+ + A^*abaA^*$.

Consider the inequality $x^2 \leq L$ and the equation $x^2 = L$. Write $\{u_1, \ldots, u_k\}$ instead of $\{u_1, \ldots, u_k\}\sim$. The elements of S with $x^2 \leq \chi(L)$ are exactly $\{aba\}$, $\{ab\}$, $\{ba\}$ (not, for instance, $\{ab, ba\}$ since $\{ab, ba\}^2 = \{bab, a, aba\} = \{1\}$). The last two are maximal solutions.

The corresponding languages are
$\psi(\{ab\}) = ab \approx \cup\ aba \approx = a^+b^2(b + a^+b^2)^*a^+b + a^+b + A^*abaA^*$,
$\psi(\{ba\}) = ba \approx \cup\ aba \approx = b(b + a^+b^2)^*a^+ + A^*abaA^*$.

In all three cases we have $x^2 = \{aba\}$ in S but not $x^2 = L$ in $\mathcal{L}$.

Consider the inequality $x^2 \leq L$ and the equation $x^2 = L$. Write $\{u_1, \dots, u_k\}$ instead of $\{u_1, \dots, u_k\}\sim$. The elements of S with $x^2 \leq \chi(L)$ are exactly $\{aba\}$, $\{ab\}$, $\{ba\}$ (not, for instance, $\{ab, ba\}$ since $\{ab, ba\}^2 = \{bab, a, aba\} = \{1\}$). The last two are maximal solutions.

The corresponding languages are
$\psi(\{ab\}) = ab \approx \cup\ aba \approx = a^+b^2(b + a^+b^2)^*a^+b + a^+b + A^*abaA^*$,
$\psi(\{ba\}) = ba \approx \cup\ aba \approx = b(b + a^+b^2)^*a^+ + A^*abaA^*$.

In all three cases we have $x^2 = \{aba\}$ in S but not $x^2 = L$ in $\mathcal{L}$.

Summarizing, $x^2 \subseteq L$ has two maximal solutions in $\mathcal{L}$ and $x^2 = L$ is not solvable there.

Universal automaton

Let $L \subseteq A^*$ be a regular language. Let

$$\mathsf{D} = \{u^{-1}L \mid u \in A^*\} = \{u_1^{-1}L, \ldots, u_n^{-1}L\} \ ,$$

Universal automaton

Let $L \subseteq A^*$ be a regular language. Let

$$\mathsf{D} = \{u^{-1}L \mid u \in A^*\} = \{u_1^{-1}L, \ldots, u_n^{-1}L\} \ ,$$
$$\widehat{\mathsf{D}} = \{Lv^{-1} \mid v \in A^*\} = \{Lv_1^{-1}, \ldots, Lv_m^{-1}\} \ ,$$

Universal automaton

Let $L \subseteq A^*$ be a regular language. Let

$$\mathsf{D} = \{u^{-1}L \mid u \in A^*\} = \{u_1^{-1}L, \ldots, u_n^{-1}L\} \;,$$
$$\widehat{\mathsf{D}} = \{Lv^{-1} \mid v \in A^*\} = \{Lv_1^{-1}, \ldots, Lv_m^{-1}\} \;,$$
$$\mathsf{U} = \{\; w_1^{-1}L \cap \cdots \cap w_k^{-1}L \mid k \geq 0,\; w_1, \ldots, w_k \in A^* \;\} \;.$$

Universal automaton

Let $L \subseteq A^*$ be a regular language. Let

$$\begin{aligned} \mathsf{D} &= \{u^{-1}L \mid u \in A^*\} = \{u_1^{-1}L, \ldots, u_n^{-1}L\} \ , \\ \widehat{\mathsf{D}} &= \{Lv^{-1} \mid v \in A^*\} = \{Lv_1^{-1}, \ldots, Lv_m^{-1}\} \ , \\ \mathsf{U} &= \{\ w_1^{-1}L \cap \cdots \cap w_k^{-1}L \mid k \geq 0,\ w_1, \ldots, w_k \in A^*\ \} \ . \end{aligned}$$

Let $B = (\beta_{ij})$ be a matrix of type m/n with entries from $\{0, 1\}$ where $\beta_{ij} = 1$ if and only if $u_j v_i \in L$. This matrix is called the **basic matrix** of the language L.

Universal automaton

Let $L \subseteq A^*$ be a regular language. Let

$$\mathsf{D} = \{u^{-1}L \mid u \in A^*\} = \{u_1^{-1}L, \ldots, u_n^{-1}L\} \ ,$$
$$\widehat{\mathsf{D}} = \{Lv^{-1} \mid v \in A^*\} = \{Lv_1^{-1}, \ldots, Lv_m^{-1}\} \ ,$$
$$\mathsf{U} = \{ \ w_1^{-1}L \cap \cdots \cap w_k^{-1}L \mid k \geq 0, \ w_1, \ldots, w_k \in A^* \ \} \ .$$

Let $B = (\beta_{ij})$ be a matrix of type m/n with entries from $\{0, 1\}$ where $\beta_{ij} = 1$ if and only if $u_j v_i \in L$. This matrix is called the **basic matrix** of the language L. Adding to the columns of B new ones which are componentwise meets of sets of columns of B $(0 \wedge 0 = 0 \wedge 1 = 1 \wedge 0 = 0, \ 1 \wedge 1 = 1)$ we get the matrix C which is called the **universal matrix** of L.

The **universal automaton** of a language L is a (non-deterministic) automaton $\mathcal{U} = (\mathsf{U}\,, A, E, I, T)$ where
$(p, a, q) \in E$ if and only if $a^{-1}p \supseteq q$,

The **universal automaton** of a language L is a (non-deterministic) automaton $\mathcal{U} = (\mathsf{U}\,, A, E, I, T)$ where
$(p, a, q) \in E$ if and only if $a^{-1}p \supseteq q$,
$q \in I$ if and only if $q \subseteq L$ and $q \in T$ if and only if $1 \in q$.

The **universal automaton** of a language L is a (non-deterministic) automaton $\mathcal{U} = (\mathsf{U}\,, A, E, I, T)$ where
$(p, a, q) \in E$ if and only if $a^{-1}p \supseteq q$,
$q \in I$ if and only if $q \subseteq L$ and $q \in T$ if and only if $1 \in q$.

Note that the states of the minimal complete deterministic automaton of L correspond to the columns of B and the states of $\mathcal{U}$ correspond to the columns of C. Moreover, we can easily compute unions and intersections of states of $\mathcal{U}$ using the matrix C.

The **universal automaton** of a language L is a (non-deterministic) automaton $\mathcal{U} = (\mathsf{U}\,, A, E, I, T)$ where
$(p, a, q) \in E$ if and only if $a^{-1}p \supseteq q$,
$q \in I$ if and only if $q \subseteq L$ and $q \in T$ if and only if $1 \in q$.

Note that the states of the minimal complete deterministic automaton of L correspond to the columns of B and the states of $\mathcal{U}$ correspond to the columns of C. Moreover, we can easily compute unions and intersections of states of $\mathcal{U}$ using the matrix C.

Proposition. (i) $\mathcal{U}$ accepts L,

The **universal automaton** of a language L is a (non-deterministic) automaton $\mathcal{U} = (\mathsf{U}\,, A, E, I, T)$ where
$(p, a, q) \in E$ if and only if $a^{-1}p \supseteq q$,
$q \in I$ if and only if $q \subseteq L$ and $q \in T$ if and only if $1 \in q$.

Note that the states of the minimal complete deterministic automaton of L correspond to the columns of B and the states of $\mathcal{U}$ correspond to the columns of C. Moreover, we can easily compute unions and intersections of states of $\mathcal{U}$ using the matrix C.

Proposition. (i) $\mathcal{U}$ accepts L,
(ii) for each non-deterministic automaton $\mathcal{V} = (V, A, G, J, W)$ accepting a subset of L, the mapping

$$\phi : q \mapsto \bigcap \{\ u^{-1}L \mid q \in J \cdot u\ \}$$

is an automaton homomorphism of $\mathcal{V}$ into $\mathcal{U}$.

is an automaton homomorphism of $\mathcal{V}$ into $\mathcal{U}$.
(iii) for each $q \in \mathsf{U}$, the automaton $(\mathsf{U}, A, E, \{q\}, T)$ accepts exactly the language q.

is an automaton homomorphism of $\mathcal{V}$ into $\mathcal{U}$.
(iii) for each $q \in \mathsf{U}$, the automaton $(\mathsf{U}, A, E, \{q\}, T)$ accepts exactly the language q.

Minimalization of NFA using the universal automaton

Let $\mathcal{U} = (\mathsf{U}, A, E, I, T)$ be the universal automaton of a regular language $L \subseteq A^*$. Each $P \subseteq \mathsf{U}$ induces a subautomaton $\mathcal{U}_P = (P, A, E_P, I \cap P, T \cap P)$ of $\mathcal{U}$ where $E_P = \{(p, a, q) \in E \mid p, q \in P\}$. Clearly, the language accepted by $\mathcal{U}_P$ is a subset of L. We formulate several conditions on a subset P of U:

is an automaton homomorphism of $\mathcal{V}$ into $\mathcal{U}$.
(iii) for each $q \in$ U, the automaton $($U$, A, E, \{q\}, T)$ accepts exactly the language q.

Minimalization of NFA using the universal automaton

Let $\mathcal{U} = ($U$, A, E, I, T)$ be the universal automaton of a regular language $L \subseteq A^*$. Each $P \subseteq$ U induces a subautomaton $\mathcal{U}_P = (P, A, E_P, I \cap P, T \cap P)$ of $\mathcal{U}$ where $E_P = \{(p, a, q) \in E \mid p, q \in P\}$. Clearly, the language accepted by $\mathcal{U}_P$ is a subset of L. We formulate several conditions on a subset P of U:

(A) the automaton $\mathcal{U}_P$ **accepts** the language L,

is an automaton homomorphism of $\mathcal{V}$ into $\mathcal{U}$.
(iii) for each $q \in$ U, the automaton $(\mathrm{U}, A, E, \{q\}, T)$ accepts exactly the language q.

Minimalization of NFA using the universal automaton

Let $\mathcal{U} = (\mathrm{U}, A, E, I, T)$ be the universal automaton of a regular language $L \subseteq A^*$. Each $P \subseteq$ U induces a subautomaton $\mathcal{U}_P = (P, A, E_P, I \cap P, T \cap P)$ of $\mathcal{U}$ where $E_P = \{(p, a, q) \in E \mid p, q \in P\}$. Clearly, the language accepted by $\mathcal{U}_P$ is a subset of L. We formulate several conditions on a subset P of U:

(A) the automaton $\mathcal{U}_P$ **accepts** the language L,

(C) $\mathcal{U}_P$ is **complete**, i.e.

$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )(\ \exists\ q \in P\ )\ q \subseteq a^{-1}p$,

(I) the **initial state is covered**, i.e. $L = \bigcup \{p \in P \mid p \subseteq L\}$,

$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )(\ \exists\ q \in P\ )\ q \subseteq a^{-1}p$,

(I) the **initial state is covered**, i.e. $L = \bigcup \{p \in P \mid p \subseteq L\}$,

(D) **closeness with respect to derivatives**, i.e.

$$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )\ a^{-1}p \in P,$$

$(\forall\, p \in P \,)(\forall\, a \in A \,)(\exists\, q \in P \,)\; q \subseteq a^{-1}p$,

(I) the **initial state is covered**, i.e. $L = \bigcup \{p \in P \mid p \subseteq L\}$,

(D) **closeness with respect to derivatives**, i.e.

$$(\forall\, p \in P \,)(\forall\, a \in A \,)\; a^{-1}p \in P,$$

(K) **(Kiel)** $(\forall\, p \in P \,)(\forall\, a \in A \,)\; a^{-1}p = \bigcup \{q \in P \mid q \subseteq a^{-1}p\}$,

$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )(\ \exists\ q \in P\ )\ q \subseteq a^{-1}p$,

(I) the **initial state is covered**, i.e. $L = \bigcup\ \{p \in P \mid p \subseteq L\}$,

(D) **closeness with respect to derivatives**, i.e.

$$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )\ a^{-1}p \in P,$$

(K) **(Kiel)** $(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )\ a^{-1}p = \bigcup\ \{q \in P \mid q \subseteq a^{-1}p\}$,

(W) **(Waterloo)** $(\ \forall\ q \in \mathsf{D}\ )\ q = \bigcup\ \{p \in P \mid p \subseteq q\}$,

$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )(\ \exists\ q \in P\ )\ q \subseteq a^{-1}p$,

(I) the **initial state is covered**, i.e. $L = \bigcup\ \{p \in P \mid p \subseteq L\}$,

(D) **closeness with respect to derivatives**, i.e.

$$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )\ a^{-1}p \in P,$$

(K) **(Kiel)** $(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )\ a^{-1}p = \bigcup\ \{q \in P \mid q \subseteq a^{-1}p\}$,

(W) **(Waterloo)** $(\ \forall\ q \in \mathsf{D}\ )\ q = \bigcup\ \{p \in P \mid p \subseteq q\}$,

(LM) the **local minimality**, i.e. $P \models (A)$ but for each $p \in P$ we have $P \setminus \{p\} \models \neg(A)$,

$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )(\ \exists\ q \in P\ )\ q \subseteq a^{-1}p$,

(I) the **initial state is covered**, i.e. $L = \bigcup \{p \in P \mid p \subseteq L\}$,

(D) **closeness with respect to derivatives**, i.e.

$$(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )\ a^{-1}p \in P,$$

(K) **(Kiel)** $(\ \forall\ p \in P\ )(\ \forall\ a \in A\ )\ a^{-1}p = \bigcup \{q \in P \mid q \subseteq a^{-1}p\}$,

(W) **(Waterloo)** $(\ \forall\ q \in \mathsf{D}\ )\ q = \bigcup \{p \in P \mid p \subseteq q\}$,

(LM) the **local minimality**, i.e. $P \models (A)$ but for each $p \in P$ we have $P \setminus \{p\} \models \neg(A)$,

(L) **(Lille)** $P_l = \{\ p \in \mathsf{D} \mid p$ is union-irreducible in $(\mathsf{D}, \subseteq)$, $p \neq \emptyset\ \}$,

$(\forall\ p \in P\ )(\forall\ a \in A\ )(\exists\ q \in P\ )\ q \subseteq a^{-1}p$,
(I) the **initial state is covered**, i.e. $L = \bigcup \{p \in P \mid p \subseteq L\}$,
(D) **closeness with respect to derivatives**, i.e.

$$(\forall\ p \in P\ )(\forall\ a \in A\ )\ a^{-1}p \in P,$$

(K) **(Kiel)** $(\forall\ p \in P\ )(\forall\ a \in A\ )\ a^{-1}p = \bigcup \{q \in P \mid q \subseteq a^{-1}p\}$,
(W) **(Waterloo)** $(\forall\ q \in \mathsf{D}\ )\ q = \bigcup \{p \in P \mid p \subseteq q\}$,
(LM) the **local minimality**, i.e. $P \models (A)$ but for each $p \in P$ we have $P \setminus \{p\} \models \neg(A)$,
(L) **(Lille)** $P_l = \{\ p \in \mathsf{D} \mid p$ is union-irreducible in $(\mathsf{D}\,, \subseteq)$, $p \neq \emptyset\ \}$,
(B) **(Brno)**

$$P_b = \{\ \bigcap \{u_j^{-1}L \mid \beta_{ij} = 1\} \mid Lv_i^{-1} \text{ union-irred. in } (\ \widehat{\mathsf{D}}\,, \subseteq)\ \}.$$

$(\forall\, p \in P\,)(\forall\, a \in A\,)(\exists\, q \in P\,)\; q \subseteq a^{-1}p$,
(I) the **initial state is covered**, i.e. $L = \bigcup \{p \in P \mid p \subseteq L\}$,
(D) **closeness with respect to derivatives**, i.e.

$$(\forall\, p \in P\,)(\forall\, a \in A\,)\; a^{-1}p \in P,$$

(K) **(Kiel)** $(\forall\, p \in P\,)(\forall\, a \in A\,)\; a^{-1}p = \bigcup \{q \in P \mid q \subseteq a^{-1}p\}$,
(W) **(Waterloo)** $(\forall\, q \in \mathsf{D}\,)\; q = \bigcup \{p \in P \mid p \subseteq q\}$,
(LM) the **local minimality**, i.e. $P \models (A)$ but for each $p \in P$ we have $P \setminus \{p\} \models \neg(A)$,
(L) **(Lille)** $P_l = \{\, p \in \mathsf{D} \mid p$ is union-irreducible in $(\mathsf{D}\,, \subseteq)$, $p \neq \emptyset\, \}$,
(B) **(Brno)**

$$P_b = \{\, \bigcap \{u_j^{-1}L \mid \beta_{ij} = 1\} \mid Lv_i^{-1} \text{ union-irred. in } (\,\widehat{\mathsf{D}}\,, \subseteq)\,\}.$$

We also put $P_0 = \{\, q \in \mathsf{D} \mid q \neq \bigcup \{r \in \mathsf{U} \mid r \subseteq q,\ r \neq q\}\,\}$.

$(\ \forall\, p \in P\ )(\ \forall\, a \in A\ )(\ \exists\, q \in P\ )\ q \subseteq a^{-1}p$,

(I) the **initial state is covered**, i.e. $L = \bigcup\, \{p \in P \mid p \subseteq L\}$,

(D) **closeness with respect to derivatives**, i.e.

$$(\ \forall\, p \in P\ )(\ \forall\, a \in A\ )\ a^{-1}p \in P,$$

(K) **(Kiel)** $(\ \forall\, p \in P\ )(\ \forall\, a \in A\ )\ a^{-1}p = \bigcup\, \{q \in P \mid q \subseteq a^{-1}p\}$,

(W) **(Waterloo)** $(\ \forall\, q \in \mathsf{D}\ )\ q = \bigcup\, \{p \in P \mid p \subseteq q\}$,

(LM) the **local minimality**, i.e. $P \models (A)$ but for each $p \in P$ we have $P \setminus \{p\} \models \neg(A)$,

(L) **(Lille)** $P_l = \{\ p \in \mathsf{D} \mid p$ is union-irreducible in $(\mathsf{D}, \subseteq)$, $p \neq \emptyset\ \}$,

(B) **(Brno)**

$$P_b = \{\ \bigcap\, \{u_j^{-1}L \mid \beta_{ij} = 1\} \mid Lv_i^{-1} \text{ union-irred. in } (\widehat{\mathsf{D}}, \subseteq)\ \}.$$

We also put $P_0 = \{\ q \in \mathsf{D} \mid q \neq \bigcup\, \{r \in \mathsf{U} \mid r \subseteq q,\ r \neq q\}\ \}$.

Theorem. The following implications between our conditions hold.

(i) (D) $\implies$ (K) & (C).

(i) (D) $\Longrightarrow$ (K) & (C).

(ii) (L) $\Longrightarrow$ (I) & (K) $\Longrightarrow$ (A) $\Longrightarrow$ (W).

(i) (D) $\implies$ (K) & (C).

(ii) (L) $\implies$ (I) & (K) $\implies$ (A) $\implies$ (W).

Moreover, (iii) P_b satisfies (A).

(i) (D) $\Longrightarrow$ (K) $\&$ (C).

(ii) (L) $\Longrightarrow$ (I) $\&$ (K) $\Longrightarrow$ (A) $\Longrightarrow$ (W).

Moreover, (iii) P_b satisfies (A).

(iv) Both P_l and P_b satisfy the condition (LM).

(i) (D) $\implies$ (K) $\&$ (C).

(ii) (L) $\implies$ (I) $\&$ (K) $\implies$ (A) $\implies$ (W).

Moreover, (iii) P_b satisfies (A).

(iv) Both P_l and P_b satisfy the condition (LM).

(v) For each $P \subseteq$ U satisfying (A) we have $P_0 \subseteq P$.

(i) (D) $\implies$ (K) & (C).

(ii) (L) $\implies$ (I) & (K) $\implies$ (A) $\implies$ (W).

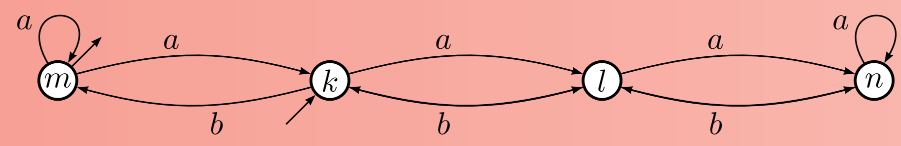
Moreover, (iii) P_b satisfies (A).

(iv) Both P_l and P_b satisfy the condition (LM).

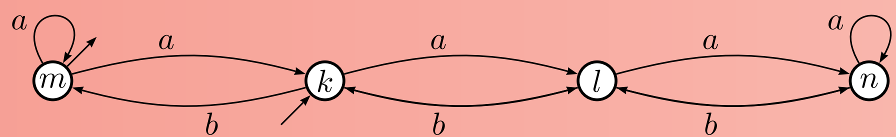
(v) For each $P \subseteq$ U satisfying (A) we have $P_0 \subseteq P$.

(vi) None of the implications in (i) and (ii) can be reversed.

Kiel Example: Consider an automaton $\mathcal{A}$:

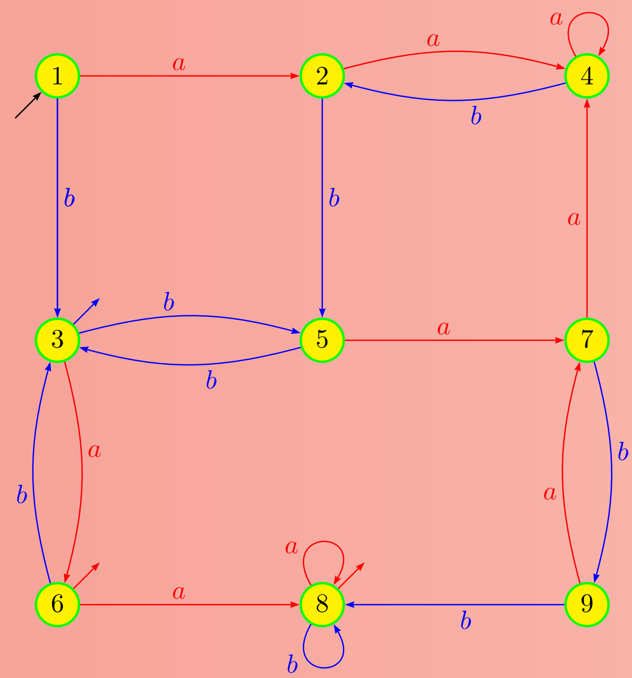


Kiel Example: Consider an automaton $\mathcal{A}$:



After determinization we get the minimal complete DFA $\mathcal{D}$:

$1 = \{k\},\ 2 = \{l\},\ 3 = \{l, m\},\ 4 = \{n\},\ 5 = \{k, n\}, \ldots$



We get the basic matrix by the determinization of $\overline{\mathcal{D}}$.

	1	2	3	4	5	6	7	8	9
1	0	0	1	0	0	1	0	1	0

We get the basic matrix by the determinization of $\overline{\mathcal{D}}$.

	1	2	3	4	5	6	7	8	9
1	0	0	1	0	0	1	0	1	0
b	1	0	0	0	1	1	0	1	1

We get the basic matrix by the determinization of $\overline{\mathcal{D}}$.

	1	2	3	4	5	6	7	8	9
1	0	0	1	0	0	1	0	1	0
b	1	0	0	0	1	1	0	1	1
b^2	0	1	1	0	0	0	1	1	1

We get the basic matrix by the determinization of $\overline{\mathcal{D}}$.

	1	2	3	4	5	6	7	8	9
1	0	0	1	0	0	1	0	1	0
b	1	0	0	0	1	1	0	1	1
b^2	0	1	1	0	0	0	1	1	1
b^3	1	0	0	1	1	1	1	1	1

We get the basic matrix by the determinization of $\overline{\mathcal{D}}$.

	1	2	3	4	5	6	7	8	9
1	0	0	1	0	0	1	0	1	0
b	1	0	0	0	1	1	0	1	1
b^2	0	1	1	0	0	0	1	1	1
b^3	1	0	0	1	1	1	1	1	1
ab^3	0	1	1	1	1	1	1	1	1

We get the basic matrix by the determinization of $\overline{\mathcal{D}}$.

	1	2	3	4	5	6	7	8	9
1	0	0	1	0	0	1	0	1	0
b	1	0	0	0	1	1	0	1	1
b^2	0	1	1	0	0	0	1	1	1
b^3	1	0	0	1	1	1	1	1	1
ab^3	0	1	1	1	1	1	1	1	1
a^2b^3	1	1	1	1	1	1	1	1	1

We get the basic matrix by the determinization of $\overline{\mathcal{D}}$.

	1	2	3	4	5	6	7	8	9
1	0	0	1	0	0	1	0	1	0
b	1	0	0	0	1	1	0	1	1
b^2	0	1	1	0	0	0	1	1	1
b^3	1	0	0	1	1	1	1	1	1
ab^3	0	1	1	1	1	1	1	1	1
a^2b^3	1	1	1	1	1	1	1	1	1
	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$		$\uparrow$			

We get the basic matrix by the determinization of $\overline{\mathcal{D}}$.

	1	2	3	4	5	6	7	8	9
1	0	0	1	0	0	1	0	1	0
b	1	0	0	0	1	1	0	1	1
b^2	0	1	1	0	0	0	1	1	1
b^3	1	0	0	1	1	1	1	1	1
ab^3	0	1	1	1	1	1	1	1	1
a^2b^3	1	1	1	1	1	1	1	1	1
	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$		$\uparrow$			

The universal matrix.

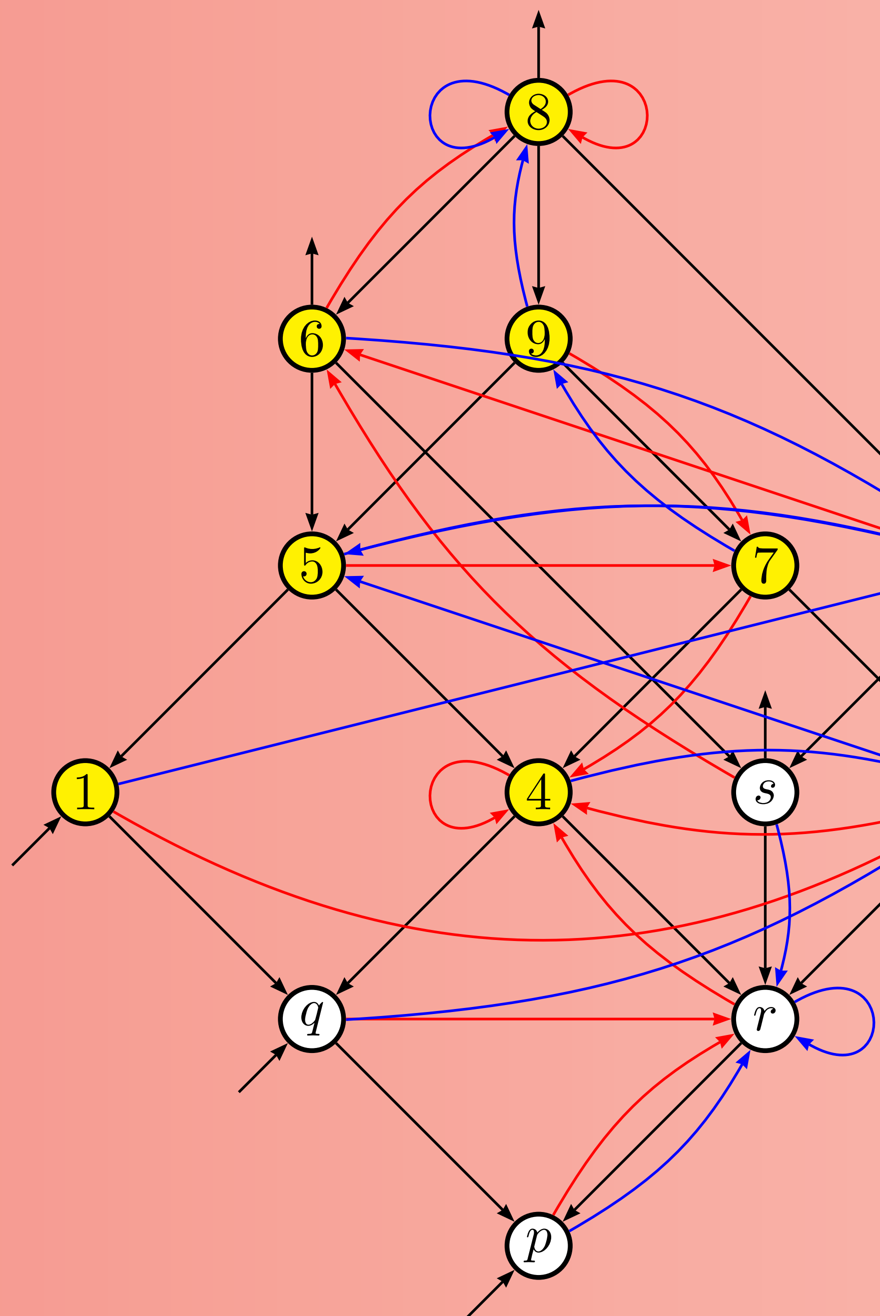
	1	2	3	4	5	6	7	8	9	p	q	r	s
1	0	0	1	0	0	1	0	1	0	0	0	0	1
b	1	0	0	0	1	1	0	1	1	0	0	0	0
b^2	0	1	1	0	0	0	1	1	1	0	0	0	0
b^3	1	0	0	1	1	1	1	1	1	0	1	0	0
ab^3	0	1	1	1	1	1	1	1	1	0	0	1	1
a^2b^3	1	1	1	1	1	1	1	1	1	1	1	1	1

		1	2	3	4	5	6	7	8	9	p	q	r	s
$\longrightarrow$	1	0	0	1	0	0	1	0	1	0	0	0	0	1
$\longrightarrow$	b	1	0	0	0	1	1	0	1	1	0	0	0	0
$\longrightarrow$	b^2	0	1	1	0	0	0	1	1	1	0	0	0	0
$\longrightarrow$	b^3	1	0	0	1	1	1	1	1	1	0	1	0	0
$\longrightarrow$	ab^3	0	1	1	1	1	1	1	1	1	0	0	1	1
	a^2b^3	1	1	1	1	1	1	1	1	1	1	1	1	1

		1	2	3	4	5	6	7	8	9	p	q	r	s
$\longrightarrow$	1	0	0	1	0	0	1	0	1	0	0	0	0	1
$\longrightarrow$	b	1	0	0	0	1	1	0	1	1	0	0	0	0
$\longrightarrow$	b^2	0	1	1	0	0	0	1	1	1	0	0	0	0
$\longrightarrow$	b^3	1	0	0	1	1	1	1	1	1	0	1	0	0
$\longrightarrow$	ab^3	0	1	1	1	1	1	1	1	1	0	0	1	1
	a^2b^3	1	1	1	1	1	1	1	1	1	1	1	1	1
		$\uparrow$	$\uparrow$									$\uparrow$	$\uparrow$	$\uparrow$

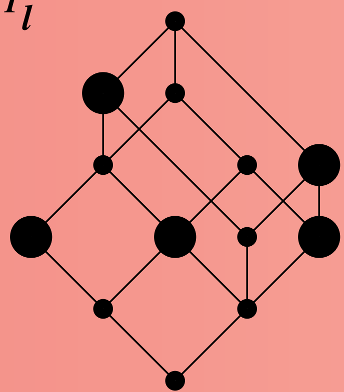
		1	2	3	4	5	6	7	8	9	p	q	r	s
$\longrightarrow$	1	0	0	1	0	0	1	0	1	0	0	0	0	1
$\longrightarrow$	b	1	0	0	0	1	1	0	1	1	0	0	0	0
$\longrightarrow$	b^2	0	1	1	0	0	0	1	1	1	0	0	0	0
$\longrightarrow$	b^3	1	0	0	1	1	1	1	1	1	0	1	0	0
$\longrightarrow$	ab^3	0	1	1	1	1	1	1	1	1	0	0	1	1
	a^2b^3	1	1	1	1	1	1	1	1	1	1	1	1	1
		$\uparrow$	$\uparrow$									$\uparrow$	$\uparrow$	$\uparrow$

The universal automaton follows.

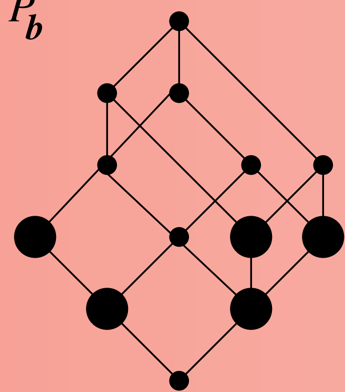

8
6
9
5
7
3
1
4
s
2
q
r
p

Consider the following choices for the set P.

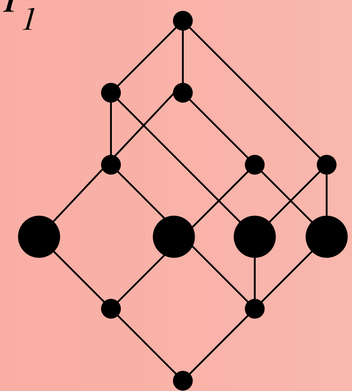
P_l



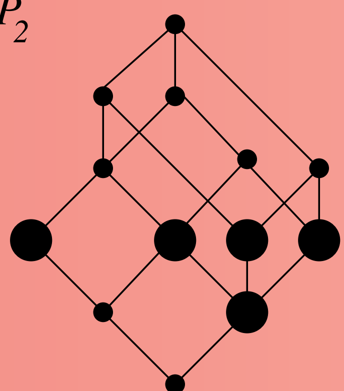
P_b



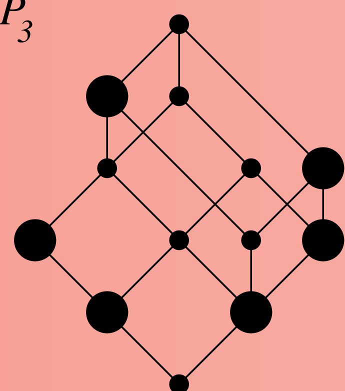
P_1



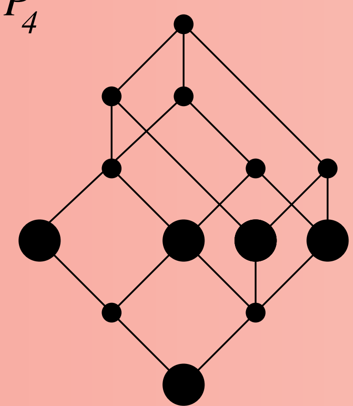
P_2



P_3



P_4



Here $P_0 = \{1, 2\}$ and the locally minimal sets of states with respect to (A) are exactly $P_l,\ P_b,\ P_1$ and P_3. Note also that P_1 is the ϕ-image of the automata we started with and P_2 is the image of its completion. Further, $P_b,\ P_2 \models (C),\ \neg(D),\ (K),\ P_1 \models \neg(C),\ \neg(D),\ \neg(K),\ P_4 \models (C),\ \neg(D),\ \neg(K)$.

Here $P_0 = \{1, 2\}$ and the locally minimal sets of states with respect to (A) are exactly $P_l,\ P_b,\ P_1$ and P_3. Note also that P_1 is the ϕ-image of the automata we started with and P_2 is the image of its completion. Further, $P_b,\ P_2 \models (C),\ \neg(D),\ (K),\ P_1 \models \neg(C),\ \neg(D),\ \neg(K),\ P_4 \models (C),\ \neg(D),\ \neg(K)$.

Waterloo Example:

The following table presents an (incomplete) DFA.

	$\downarrow$ 1	2	3	4	$\uparrow$ 5	$\uparrow$ 6	7	8
a	5	6	7	3	3	2	1	1
b	–	–	2	8	8	4	4	–

Here $P_0 = \{1, 2\}$ and the locally minimal sets of states with respect to (A) are exactly P_l, P_b, P_1 and P_3. Note also that P_1 is the ϕ-image of the automata we started with and P_2 is the image of its completion. Further, P_b, $P_2 \models (C)$, $\neg(D)$, (K), $P_1 \models \neg(C)$, $\neg(D)$, $\neg(K)$, $P_4 \models (C)$, $\neg(D)$, $\neg(K)$.

Waterloo Example:

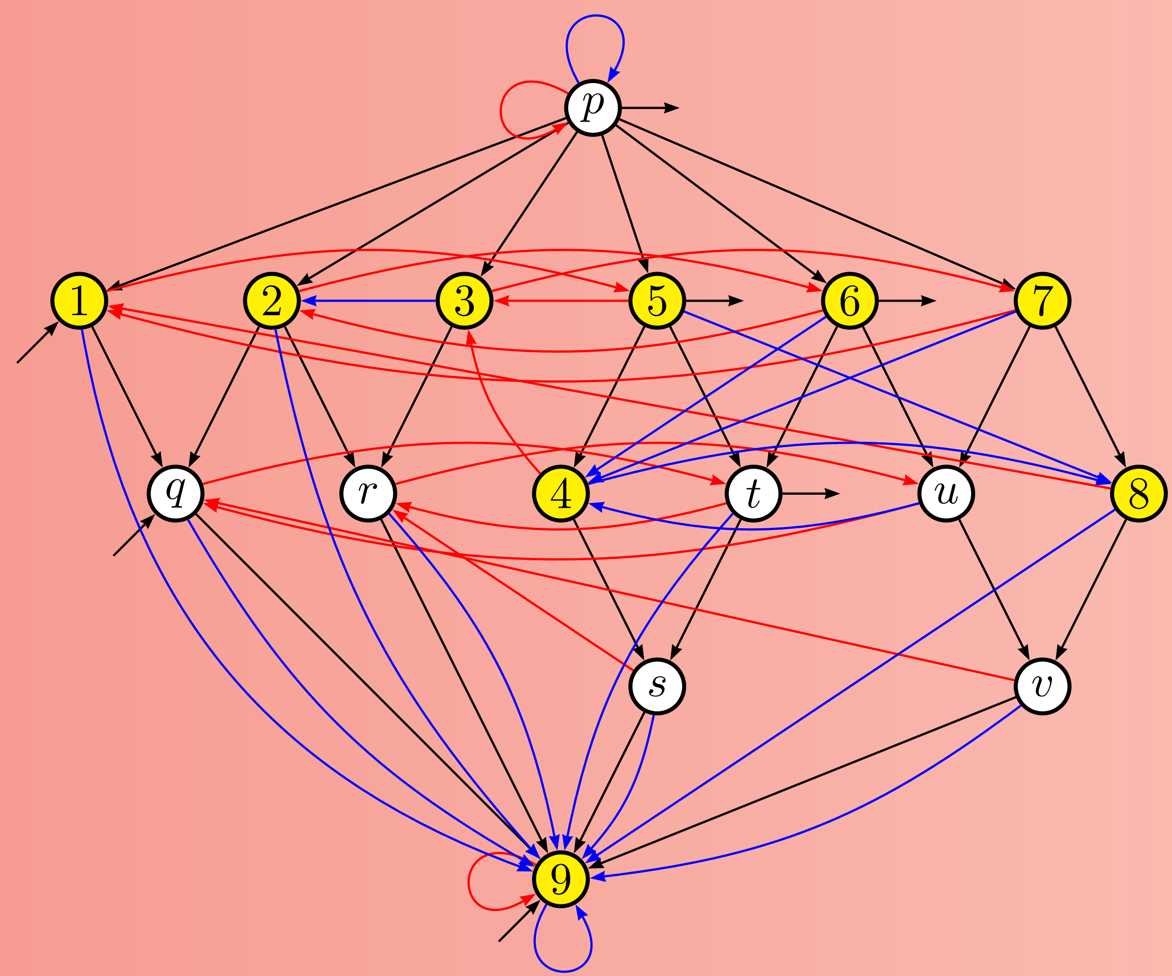
The following table presents an (incomplete) DFA.

	$\downarrow$ 1	2	3	4	$\uparrow$ 5	$\uparrow$ 6	7	8
a	5	6	7	3	3	2	1	1
b	–	–	2	8	8	4	4	–

Its universal matrix follows.

	1	2	3	4	5	6	7	8	9	p	q	r	s	t	u	v
1	0	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0
a	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0
b	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
a^2	0	0	0	0	0	1	1	1	0	1	0	0	0	0	1	1
ba	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
a^3	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	0
aba	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0
a^4	0	0	0	1	1	1	0	0	0	1	0	0	1	1	0	0
a^2ba	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
a^3ba	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0
ba^4	0	0	0	0	0	1	1	0	0	1	0	0	0	0	1	0

The universal automaton is depicted below.



We have $P_0 = \{1, 3, 4, 8\}$ and the locally minimal sets of states with respect to (A) are exactly $P_l = P_0 \cup \{2, 5, 6, 7\}$, $P_b = P_0 \cup \{q, r, t, u\}$, $P_1 = P_0 \cup \{2, 5, 6, u\}$, $P_2 = P_0 \cup \{2, 6, t, u\}$ and $P_3 = P_0 \cup \{2, 6, 7, t\}$. We have that $P_4 = P_0 \cup \{2, t, u\} \models (W), \neg(A)$.